Evasion of Guilt in Expert Advice*

Kiryl Khalmetski†

June 6, 2019

Abstract

We develop a model of strategic communication between an uninformed receiver and a partially informed sender who is guilt-averse toward the receiver. The sender's cost of sending a particular message is endogenous, depending on the receiver's *beliefs* induced by this message rather than on its exogenous formulation. Such preferences lead to the endogenous emergence of evasive communication in that the sender types who prefer not to reveal their information to the receiver pool with uninformed types rather than with types observing different information. As a result, the receiver may prefer an equilibrium with a smaller amount of messages used on the equilibrium path. Besides, dealing with an ex ante less informed sender can be beneficial to the receiver, while the sender himself may want to commit to a smaller ex ante likelihood of being informed.

Keywords: guilt aversion, information transmission, experts, psychological game theory.

JEL codes: D82, D83, D84, C72.

1 Introduction

Many settings involve monetary incentives for strategic misrepresentation of information transmitted from an informed expert (e.g., a financial advisor or a doctor) to an uninformed customer (Anagol et al. 2017, Johnson and Rehavi 2016). Still, even in the presence of such conflict of interest, customers considerably rely on this service in practice.[1] Thereby, they rely also on the indirect costs arising for the expert from

*Accepted for publication in *Journal of Economic Behavior & Organization*.

†University of Cologne, Department of Economics, Albertus-Magnus-Platz, 50923 Cologne, Germany (e-mail: kiryl.khalmetski@uni-koeln.de). I am thankful to Matthias Blonski, Ruben Enikolopov, Guido Friebel, Andreas Hackethal, Roman Inderst, Michael Kosfeld, Axel Ockenfels and Ferdinand von Siemens for helpful comments. Financial support of the German Research Foundation (DFG) through the Research Unit "Design & Behavior" (FOR 1371) is gratefully acknowledged.

[1]For example, a large online survey by Chater et al. (2010) shows that nearly 58 percent of purchasers of investment products are influenced by advisors.

deceiving the customer, which may limit the scope of fraudulent advice. For example, deception can lead to reputational loss (Bolton et al. 2007), reclamation costs (Inderst and Ottaviani 2013), or psychological costs which arise from intrinsic concern for the well-being of the other party (McGuire 2000, Kesternich et al. 2015).[2] The present paper theoretically examines the role of another behavioral motivation that enhances credibility of communication under monetary conflict of interest - namely, guilt aversion, i.e. a preference to comply with the expectations of others (Battigalli and Dufwenberg 2007).[3]

In sum, we study the implications of guilt aversion for the structure of equilibrium (cheap-talk) communication between a sender and a receiver under monetary conflict of interest. One of the key elements of the model is that the sender is uninformed with some probability. This allows the sender to engage in strategic evasion, i.e. to (credibly) pretend to be uninformed about the state of the world. We show that guilt aversion on the side of the sender leads to the endogenous emergence of such evasive communication in equilibrium. In turn, this has consequences for both players' welfare as well as for their preferences regarding the ex ante quality of the sender's information.

Specifically, in our model setting the sender observes the state of the world, which can be either good or bad, with a certain probability, while with the remaining probability he is uninformed. Then, the sender sends a message to the receiver out of an arbitrarily large (countable) message space, or refrains from advice. Finally, the receiver must decide between a riskless action (abstaining) and a risky action (investment), with the latter having a positive payoff for her only in the good state.

The sender is biased to always induce investment independently from the state of the world, while at the same time being sensitive to guilt toward the receiver (in the main specification of the model). Guilt is determined by the discrepancy between the receiver's payoff expectation conditional on the sender's message and the ex post receiver's payoff. The sender's guilt sensitivity is unobservable to the receiver. Thus, the sender is characterized by both the information that he has observed and his guilt sensitivity (the latter referred to as the sender's type).

There are only two qualitatively different equilibria in this game (while all other existing equilibria are payoff-equivalent to either of them). In the *pooling* equilibrium, the sender types who induce investment send the same message, while all other types refrain from advice. In the *separating* equilibrium, the sender types who induce investment after observing the bad state of the world separate from the types observing the good state, and send instead an "evasive" message pooling (exclusively) with uninformed types. The evasive message leads to lower receiver's beliefs than the message used in the pooling

[2] See Gneezy (2005), Sutter (2009), Lundquist et al. (2009), Erat and Gneezy (2012) and Fischbacher and Föllmi-Heusi (2013) for experimental evidence for aversion to lying.

[3] Guilt aversion has gained significant empirical support in recent years. Experimental evidence is documented in Guerra and Zizzo (2004), Charness and Dufwenberg (2006), Reuben et al. (2009), Khalmetski et al. (2015), Khalmetski (2016) and Ederer and Stremitzer (2017), among others. In particular, Charness and Dufwenberg (2006), Khalmetski (2016) and Ederer and Stremitzer (2017) show the relevance of guilt aversion in the context of bilateral communication. Note that Ellingsen et al. (2010) and Vanberg (2008) do not find support for guilt aversion; see Khalmetski et al. (2015) for a possible reconciliation.

equilibrium, and hence is less costly for the sender in terms of expected guilt. Notably, the separating equilibrium does not arise under purely outcome-based preferences, which lead to the pooling equilibrium (while all other existing equilibria are payoff-equivalent).

A shift from the pooling to the separating equilibrium can be both beneficial and detrimental to the receiver. In particular, the separating equilibrium provides the sender with a psychologically cheap way to induce investment in the bad state of the world by sending the evasive message. This tends to increase the rate of unprofitable investment in the bad state. At the same time, in the separating equilibrium the truly uninformed sender types have lower expected guilt from inducing (ex ante profitable) investment, which leads to more efficient receiver's investment choices. This positive effect dominates whenever the monetary conflict of interest between the sender and the receiver is sufficiently small.

We also consider the players' preferences over the ex ante quality of the sender's private information. One of the results is that dealing with an (ex ante) *less* informed sender can be preferable for the receiver. This occurs due to the fact that the ex ante probability that the sender is uninformed affects the receiver's beliefs conditional on the message, and hence the expected guilt of the sender. As a result, an (ex ante) more knowledgeable sender can be, at the same time, more prone to inducing receiver's response which is suboptimal for the latter. Under certain parameter values, this effect is sufficiently strong to outweigh the positive effect of higher quality of the sender's private information. Similarly, the sender himself may want to ex ante commit to a smaller likelihood of being informed, since this may reduce the receiver's expectations conditional on his message, and hence result in smaller guilt. Again, these effects cannot be obtained under outcome-based preferences, in which case both players are never worse off if the sender is more likely to be informed.

Our study relates to several strands of literature. Kartik et al. (2007) and Kartik (2009) study message-based cost of lying, which depends only on how much the exogenously given formulation of a message quantitatively deviates from the truth. While such approach can address a broad range of situations (like reporting of company profits to shareholders), there are limits to its applicability. First, it may occur that the states of the world which are to be reported cannot be ranked quantitatively (e.g., possible diagnoses of a patient), so that different possible lies cannot be compared by severity based only on message formulations. Second, there are many ways in which the expert can manipulate or mitigate explicit message formulations while conveying the same meaning (e.g., euphemisms). In contrast, our approach based on guilt aversion provides a measure of the cost of communication which can be applied in both of these cases: the difference between expectations induced by the advice and the actually realized outcome.[4]

Khalmetski et al. (2017) consider a structurally similar setting (both theoretically and experimentally) where a sender, who may be either informed or uninformed about a binary state of the world, sends a message to a risk averse receiver. They also distinguish separating and pooling communication strategies (terming them as "evasive" and "direct" lying, respectively), showing that senders may use evasion (i.e. pretending

[4]See Sobel (2018) for an analysis and discussion of belief-driven vs. message-driven measures of distortion in the sender's communication in sender-receiver games.

to be uninformed) to sidestep higher psychological costs from outright lying. At the same time, their model uses a reduced-form approach by exogenously assuming that the (fixed) intrinsic cost of evasive lying is lower than the cost of direct lying. In the current setting, this basic feature of the model is derived endogenously, which allows for further theoretical implications.

The role of guilt aversion in communication was studied by Charness and Dufwenberg (2006) and Ederer and Stremitzer (2017). However, in their settings, communication (in particular, giving promises to behave in a certain way) serves effectively as a commitment device for a guilt-averse agent in games involving moral hazard. The current setting is different in that communication resolves information asymmetry between the sender and the receiver.

More relatedly, Loginova (2012) and Battigalli et al.(2013) consider communication of private information with a guilt-averse sender. Loginova (2012) studies implications of guilt aversion for the cheap-talk setting of Crawford and Sobel (1982), and finds that higher guilt aversion of the sender allows for more informative communication. Battigalli et al.(2013) show that guilt aversion can organize the experimental data in Gneezy (2005), in particular, predicting that the sender is less likely to lie the larger is the discrepancy in the receiver's payoffs between the possible outcomes. At the same time, in both of these studies, the sender is always informed about the state of the world, so that there is no scope for strategic pooling with uninformed types as a means to mitigate guilt, which plays a central role in our model.[5]

The problem of strategic evasion has been analyzed thus far mainly in verifiable disclosure settings (Dye 1985, Dziuda 2011, Bhattacharya et al. 2018), where the sender cannot misreport the observed information, but can only conceal it. Austen-Smith (1994) studies evasive communication in a mixed setting, where an informed sender can choose any message, while the uninformed sender cannot conceal the fact that he is uninformed from the receiver. In contrast to these studies, we show that, once the sender has belief-dependent preferences, a credible evasion can emerge even with completely unrestricted communication.

The remainder of this paper is organized as follows. Section 2 presents the model. Section 3 analyzes existing equilibria. Section 4 studies the effect of the sender's ex ante information quality on the welfare of both players. Section 5 presents final discussion. All proofs are in the online Appendix.

[5] A different type of belief-dependent preferences in the context of communication games - concern for being believed to tell the truth - has been modeled in Abeler et al. (forthcoming), Dufwenberg and Dufwenberg (2018), Gneezy et al. (2018) and Khalmetski and Sliwka (forthcoming).

	State G	State B
$a^s = \varnothing$, $a^r = I$	$0, P$	$0, c$
$a^s = \varnothing$, $a^r = A$	$0, 0$	$0, 0$
$a^s \in M$, $a^r = I$	F, P	F, c
$a^s \in M$, $a^r = A$	$0, 0$	$0, 0$

Figure 1: Payoff matrix of Investment Game.

2 The model

2.1 Investment Game

We consider the following *Investment Game*. The game is between two players, the sender (he) and the receiver (she). There are two possible states of the world $\tau \in \{G, B\}$ (*good* and *bad*, respectively), each occurring with prior probability 1/2. The state of the world is privately observed by the sender with probability $\kappa \in (0, 1)$. That is, there are three possible states of sender information $i^s \in \{G', B', N'\}$ (termed below as information states), where G' corresponds to observation of G, B' to observation of B, and N' to no information.[6]

The timing of the game is as follows. In stage 1 the sender chooses action a^s,[7] which can be either of the following:

- sending a message m to the receiver about the state of the world out of a countable message space M with $|M| \geq 2$ (we impose no other structure on the message space, i.e. the exogenous formulation of the messages is completely irrelevant),

- refraining from giving advice (which is denoted by $\varnothing$).

In stage 2, the receiver takes a binary action $a^r \in \{I, A\}$ (*invest* or *abstain*, respectively) and the payoffs are realized.

The monetary payoffs are as follows. If the sender refrains from giving advice, his payoff is always 0 independently of the receiver's action and the state of the world. If the sender sends a message, then he gets F if the receiver invests following his advice (independently of the state of the world), and 0 otherwise. The receiver gets 0 if she abstains from investment, and a state-contingent payoff $\pi(\tau)$ in case of investment such that $\pi(G) = P$ and $\pi(B) = c$. The payoff matrix is given in Figure 1.

Regarding the payoffs, we assume:

1) $F > 0$: the sender prefers investment independently of the state of the world (in case of sending a message);

2) $P > 0$, $c < 0$: the receiver prefers to invest only in the good state.

Thus, there is monetary conflict of interest between the sender and the receiver in the bad state of the world. In terms of applications, a financial advisor can be, for instance, monetarily biased toward recommending investment in a specific financial product, which allows him to receive a higher commission (independently of whether this product fits

[6] The main qualitative results are robust to adding more information states, e.g. if the sender (besides being either fully informed or uninformed) can also obtain a noisy signal about the state of the world.

[7] Hereinafter, the upper index r refers to the receiver and s to the sender.

the receiver's needs). In a similar way, a doctor can be incentivized by a pharmaceutical company to prescribe its products to patients.

We also make the following additional assumption about the payoffs.

Assumption 1 *Investment is ex ante profitable for the receiver, i.e.* $P > -c$.

This restriction is necessary to generate evasive communication in equilibrium (considered in Section 3.3.3). Otherwise, evasion, i.e. mimicking the uninformed types, cannot induce investment.[8]

The assumption $0 < \kappa < 1$ is also crucial for our setting (to ensure the credibility of evasive communication), and reflects the fact that the sender might sometimes fail to adequately address the receiver's investment problem (while being aware of this fact). For example, a doctor might not always be able to detect the true cause of a patient's symptoms (and hence, to recommend the right medical treatment), due to the complexity of the patient's case, lack of specific experience or competence, or noisy information from diagnostic tests.

2.2 Preferences

2.2.1 Receiver's preferences

Denote by $u^r(a^r, \tau)$ the ex post utility of the receiver from action a^r in state of the world τ. Her expected utility conditional on the sender's action a^s and investment is:

$$E^r[u^r(I, \tau)|a^s] = \eta(a^s)P + (1 - \eta(a^s))c, \tag{1}$$

where $\eta(a^s) \equiv \Pr^r[G|a^s]$ is the receiver's belief about the state of the world conditional on the sender's action a^s.[9] If the sender sends message m, we call $\eta(m)$ the persuasiveness of the message (in the sense of how persuasive is the message in inducing investment). The receiver's utility from abstaining, $u^r(A)$, is always 0 (see Figure 1). Thus, she prefers to invest if and only if her expected utility from investment is larger than the utility from abstaining, i.e.[10]

$$\begin{aligned} E^r[u^r(I, \tau)|a^s] &> 0 \\ \Leftrightarrow \quad \eta(a^s) &> \frac{-c}{P - c} \equiv \underline{\eta}. \end{aligned} \tag{2}$$

2.2.2 Sender's preferences

Outcome-based preferences. As a benchmark case, we consider purely outcome-based preferences of the sender, i.e. when he cares solely about his and the receiver's monetary

[8] In this case, the only equilibrium is the pooling equilibrium, considered in Section 3.3.2.

[9] Although formally this specification corresponds to risk neutrality, risk aversion does not qualitatively change any of the subsequent results as far as investment is still ex ante profitable.

[10] Thus, we assume that the receiver prefers abstaining over investment conditional on equal utility, which is (essentially) without loss of generality for the subsequent results as far as only pure strategies are considered.

outcomes.[11] In particular, assume that the sender has some fixed cost ρ of incurring the receiver's losses, e.g., arising from inequality aversion or potential legal liability, which he bears if the receiver gets a negative payoff of c.[12] Then, the sender's expected utility conditional on investment can be represented in the following form:

$$U^s_{i^s}(\theta, I) = F - \Pr[B|i^s] \cdot \theta\rho, \tag{3}$$

where θ is a sensitivity parameter, which is unknown to the receiver and uniformly distributed on $(0, \bar{\theta}]$ (for consistency with the model of guilt aversion considered below). The sender's utility conditional on the receiver's abstaining is 0 for all sender types, since then the receiver does not incur losses. Besides, the sender is not deemed accountable for the receiver's losses if he refrains from advice, in which case the sender's utility is equal to his monetary payoff of 0 independently of the receiver's action.

Guilt aversion. In the main specification of the model, the sender is assumed to be guilt-averse, i.e. he dislikes to be responsible for disappointing the receiver's expectations (Battigalli and Dufwenberg 2007).[13] Such disappointment arises if the sender's message induces overly high expectations relative to the eventually realized outcome. Specifically, the receiver's expectations induced by message m are disappointed whenever her expected utility (conditional on m) is higher than her ex post payoff:

$$D^r(m, a^r, \tau) = \max\left\{0, E^r[u^r(a^r, \tau)|m] - u^r(a^r, \tau)\right\}, \tag{4}$$

where $D^r(m, a^r, \tau)$ is the magnitude of disappointment.

Next, we assume that in case of sending message m, the expected guilt of the sender in information state i^s is

$$G^s_{i^s}(\theta, m, a^r) = \theta E^s[D^r(m, a^r, \tau)|i^s], \tag{5}$$

where θ is the sender's sensitivity toward guilt. Finally, the total expected utility of the sender in information state i^s in case of sending message m, denoted by $U^s_{i^s}(\theta, m, a^r)$, is assumed to be additive in the monetary and guilt components:

$$U^s_{i^s}(\theta, m, a^r) = F \cdot 1_I - G^s_{i^s}(\theta, m, a^r), \tag{6}$$

where 1_I is an indicator function, equal to 1 if the receiver invests and zero otherwise.

Note that $D^r(m, a^r, G) = 0$ since the receiver cannot be disappointed by the highest

[11] This subsumes purely selfish preferences, when the sender does not care about the receiver's payoff, as a special case. Under such preferences, the sender would always prefer to induce investment independently of the state of the world.

[12] For simplicity, we assume that the sender has no change in utility if the receiver gets P, although the equilibrium predictions remain qualitatively the same if the sender would be similarly affected also in this case. The only assumption that matters here is that the sender's utility does not directly depend on beliefs.

[13] This concept originates from psychological game theory, which presumes that utility can depend on beliefs *per se* (Geanakoplos et al. 1989, Battigalli and Dufwenberg 2009).

possible outcome in the good state (i.e. $E^r[u^r(a^r, \tau)|m] \leq P$). Consequently, the sender expects non-zero guilt after sending message m and the receiver investing if and only if the bad state of the world is realized, which implies

$$U^s_{i^s}(\theta, m, I) = F - \theta\lambda_{i^s}(E^s E^r[u^r(a^r, \tau)|m] - c), \tag{7}$$

where λ_{i^s} is the sender's probability of the bad state conditional on i^s. If the receiver abstains conditional on the message, then her outcome is no longer stochastic (being zero in all states), so that $E^r[u^r(A)|m] = u^r(A) = 0$, implying

$$D^r(m, A, \tau) = G^s_{i^s}(\theta, m, A) = U^s_{i^s}(\theta, m, A) = 0. \tag{8}$$

Hence, the sender never expects guilt if the receiver abstains following his message.

As in the case of the outcome-based preferences, the sender is not deemed (either psychologically or legally) responsible for the receiver's losses if he refrains from advice, in which case the sender's utility is 0 independently of the receiver's action. Thus, the sender always has an opportunity to completely avoid guilt by just refraining to provide any advice to the receiver.

The sender's guilt aversion coefficient θ is a random variable, unknown to the receiver, distributed uniformly on an interval $(0, \bar{\theta}]$. This assumption serves to reflect the uncertainty of the receiver about the trustworthiness of the sender, which is widely heterogeneous in the population as documented by many experimental studies (e.g., Charness and Dufwenberg 2006).[14] Hence, the sender is characterized by both the information which he has observed i^s and his sensitivity to guilt θ. In what follows, we refer to θ as the sender's "type".

Note that the communication costs defined in the current model may relate not only to psychological guilt. The term $D^r(m, a^r, \tau)$, more generally, is supposed to reflect the receiver's dissatisfaction with advice arising from her frustration from unfulfilled expectations (induced by this advice). In turn, the receiver's dissatisfaction with advice can naturally lead to other costly consequences for the sender besides psychological costs, for example, reputational losses. Naturally, our definition still leaves aside some other aspects of communication costs, like aversion to induce a wrong decision of the receiver with own advice (due to, for instance, altruistic concerns for the receiver). In particular, by (8) the sender bears no costs if the receiver abstains, even if the sender knows that it is a suboptimal decision. At the same time, it is plausible to assume that in many settings the advisor is disciplined by the receiver's *perception* of the quality of advice rather than its actual quality.[15] These are the settings which fit our model and which are thus at the

[14] From a theoretical perspective, a sufficiently wide distribution of θ makes the endogenous variables of the model (e.g. the frequency of investment in a given information state) continuously depend on the exogenous parameters, which in turn allows for richer implications in terms of comparative statics.

[15] In particular, it is plausible that in the corresponding real-life settings the receiver might not find out the actual state of the world if she abstains from investment. For instance, she might never realize whether some innovative product fits her preferences unless she really tries the product.

focus of the current analysis.[16]

Lexicographic preferences. The above specification of the receiver's preferences implies that the sender is always certain of the receiver's investment conditional on m in equilibrium if $\eta(m) > \underline{\eta}$ (in which case the receiver prefers investment over abstaining). However, one can think of a more general setting where the receiver's likelihood of investment varies more smoothly with respect to her beliefs conditional on m. This would be the case, for instance, if the sender were unaware of the receiver's exact degree of risk aversion, which could take values on a sufficiently wide range. In such a case, the likelihood of investment from the perspective of the sender would increase with $\eta(m)$ (also if $\eta(m)$ goes beyond $\underline{\eta}$). Then, a sender type preferring receiver's investment over abstaining would, all else equal, strictly prefer the message leading to the highest likelihood of investment. While we omit a full modeling of this aspect for the sake of expositional simplicity,[17] it is reasonable to leave it in the model at least in the form of lexicographic preferences.

Assumption 2 *If two messages lead to investment and yield the same positive (negative) expected utility for the sender, then he strictly prefers the message inducing a higher (lower) receiver's belief.*

Note that under guilt aversion this assumption can effectively apply only to types in state G' (who are certain that the state of the world is good, and hence experience no guilt), while types in states B' and N' will always strictly prefer one of several messages if these messages induce investment while leading to different receiver's beliefs (and hence, different levels of expected guilt).

Besides, it is natural to assume that conditional on equal expected utility, the sender prefers sending a message over completely refraining from advice (e.g., due to some reputational concerns for being perceived as knowledgeable).

Assumption 3 *If a message leads to the expected utility of 0, the sender strictly prefers this message over refraining from advice.*

2.3 Solution concept

The equilibrium outcome is characterized by

1. the strategy of the receiver $\sigma^r : \{M, \varnothing\} \to \{I, A\}$ specifying whether to invest or abstain conditional on each possible message and sender's refrainment from advice;

[16] Note that we do not model any costs of lying stemming from the exogenous formulation of the messages (i.e. their "literal" meaning). First, adding this additional layer of preferences would not change the qualitative predictions of the model. Second, the main aim of our analysis is to disentangle the effects stemming purely from belief-driven communication costs, which motivates having such costs as the only source of informativeness of communication in our model.

[17] See Khalmetski et al. (2017) where this property of the receiver's investment is explicitly incorporated into the analysis in a similar setting.

2. the strategy of the sender $\sigma^s : (0, \bar{\theta}] \times i^s \rightarrow \{M, \varnothing\}$ specifying whether to send a particular message or refrain from advice for each sender type θ and information state i^s;

3. the receiver's belief about the state of the world conditional on each sender's action $\eta(a^s)$;[18]

4. all higher-order beliefs about the state of the world conditional on each sender's action (for the analysis of guilt aversion).

We apply the solution concept of pure-strategy perfect Bayesian equilibrium, which implies that the sender's and receiver's equilibrium strategies should maximize the respective expected utility functions given equilibrium beliefs; the receiver's first-order beliefs are derived by Bayes' rule whenever possible; higher-order beliefs are correct (Battigalli and Dufwenberg 2009).

We define equilibria as *payoff-equivalent* if the receiver and each sender type in each information state have the same expected payoff (including the psychological payoff) between these equilibria. We define an equilibrium where the receiver always invests (conditional on any sender's action on the equilibrium path) as *unresponsive*, and otherwise *responsive*.

3 Equilibrium analysis

3.1 General properties

Let us consider general properties of all existing equilibria (under either guilt aversion or outcome-based preferences). First, let us observe that at least one message must be sent in any equilibrium.

Lemma 1 *There exists no equilibrium where all sender types refrain from advice.*

The reason for the result is that if the sender always refrains from advice, then for any possible out-of-equilibrium beliefs at least some sender types would have a strict incentive to deviate to an out-of-equilibrium message.

Based on this, one can show the following result.

Lemma 2 *In any equilibrium, there exists at least one message leading to investment.*

The intuition is that if at least one message is sent in equilibrium (which always holds by the previous lemma), then at least one of them must induce beliefs (i.e. the probability of the good state conditional on the message) not lower than 0.5 — otherwise, there is a

[18] The receiver can also form beliefs about the sender's type θ and his information state i^s conditional on m. At the same time, specifying these beliefs in addition to $\eta(a^s)$ is redundant since the latter belief is already sufficient to determine the equilibrium strategies of both players.

contradiction to the prior of 0.5. Then, the receiver should invest after this message by Assumption 1.

Next, those types who indeed observe the good state always induce investment.

Lemma 3 *If $i^s = G'$, then all sender types send a message leading to investment.*

Indeed, if the sender observes the good state, his anticipated cost from inducing investment is zero, since he knows that the investment will turn out successful. Hence, his expected utility from any message leading to investment is F, which is strictly larger than the zero utility from sending a message leading to abstaining or refraining from advice. Finally, in any equilibrium, there is a possibility for the sender to send an investment-inducing message by Lemma 2.

In contrast to this case, whenever the sender does not observe the good state with certainty (i.e. $i^s \neq G'$), the probability assigned by him to the bad state, and hence the expected cost from inducing investment, is strictly positive. One can show that the strategy of such types has a cutoff structure in any equilibrium.

Lemma 4 *For each $i^s \in \{B', N'\}$ there exists a cutoff $\widehat{\theta}_{i^s} \in (0, \bar{\theta}]$ such that all types with $\theta < \widehat{\theta}_{i^s}$ send a message leading to investment and all types with $\theta > \widehat{\theta}_{i^s}$ (if any) send a message leading to abstaining or refrain from advice.*

The intuition behind this result is the following. First, there are always types in any information state who are sufficiently insensitive to guilt (under guilt aversion) or receiver's loss (under outcome-based preferences) to prefer inducing investment for any receiver's beliefs (which they can do by Lemma 2). Second, if some type prefers to induce investment over getting zero, then all less sensitive types would also prefer at least the same message over zero, hence, also inducing investment. Analogously, once some type prefers the utility of zero over any possible investment-inducing message, all higher types would also prefer zero, and hence would either induce abstaining or refrain from advice. This corroborates the cutoff structure described in Lemma 4.[19]

Note that Lemma 4 allows that the receiver may invest in equilibrium not only if the matched sender type θ is below the cutoff, but also if θ is above the cutoff while the sender refrains from advice (since the receiver's belief conditional on refrainment may be sufficiently high in equilibrium).

Finally, in any equilibrium the receiver's beliefs are determined by the sender's strategy through Bayes' rule that results in the following expression.

Lemma 5 *The receiver's equilibrium belief conditional on the sender's action a^s is*

$$\eta(a^s) \equiv \Pr[G|a^s] = \frac{\Pr[a^s|G']\kappa + \Pr[a^s|N'](1-\kappa)}{(\Pr[a^s|G'] + \Pr[a^s|B'])\kappa + 2\Pr[a^s|N'](1-\kappa)}. \tag{9}$$

[19] The strategy of the cutoff type $\theta = \widehat{\theta}_{i^s}$ is not specified by Lemma 4 since this type can be indifferent over inducing investment or getting 0, and hence his equilibrium strategy can be both. However, as this type has zero measure, his equilibrium strategy does not matter for the subsequent results.

Note that $\Pr[a^s|i^s]$ is determined by the sender's strategy in information state i^s, i.e. by the fraction of types who choose action a^s conditional on this information state, while κ denotes the prior probability that the sender is informed.

3.2 Outcome-based preferences

Let us consider the equilibrium characterization in the benchmark case of purely outcome-based preferences of the sender. We define a *pooling equilibrium* as follows:

Definition 1 *A pooling equilibrium is defined as an equilibrium in which*

- *Some message $\bar{m}$ is sent if $\theta \in (0, \bar{\theta}]$ in state G', $\theta \in (0, \widehat{\theta}^P_{B'}]$ in state B', and $\theta \in (0, \widehat{\theta}^P_{N'}]$ in state N';*
- *All other types in states B' and N' (if any) refrain from advice;*
- *The receiver invests after $\bar{m}$, and invests after refrainment if and only if $\eta(\varnothing) > \underline{\eta}$;*
- *The beliefs after $\bar{m}$ and $\varnothing$ are determined by Bayes' rule if possible. For any out-of-equilibrium message $\widehat{m}$ it holds that $\eta(\widehat{m}) = \eta(\bar{m})$ while the receiver invests after it.*[20]

The main feature of this equilibrium is that all types who prefer to induce investment send the same message $\bar{m}$. Thus, there is a complete pooling of investment-inducing types in states B' and N' with types in state G'.

Let us call a pooling equilibrium *essentially unique* if all other existing pooling equilibria are the same except for the exogenous value of message $\bar{m}$. Then, the following holds.

Proposition 1 *Under outcome-based preferences, there exists an essentially unique pooling equilibrium.*

The intuition for this equilibrium is the following. The sender's preferences specified by (3) imply that in any equilibrium all types with $\theta \le \widehat{\theta}_{B'} = F/\rho$ ($\theta \le \widehat{\theta}_{N'} = 2F/\rho$) prefer to induce investment in state B' (N'), while the remaining types prefer to get 0 (e.g., by refraining from advice). In turn, the receiver always prefers to invest after obtaining $\bar{m}$, i.e. $\eta(\bar{m}) > \underline{\eta}$ (see (2)). This holds due to the fact that all types in state G' pool on $\bar{m}$, which ensures $\eta(\bar{m}) \ge 0.5$ (while $\underline{\eta} < 0.5$ by Assumption 1).

Figure 2 shows the basic structure of the pooling equilibrium. Here, each horizontal line represents the set of sender types for a given information state. The black bracket indicates types who send message $\bar{m}$, while the white bracket indicates types who refrain from advice. The figure shows three possible subtypes of this equilibrium depending on

[20] Note that the definition does not restrict (out-of-equilibrium) beliefs conditional on refrainment if the latter is not on the equilibrium path.

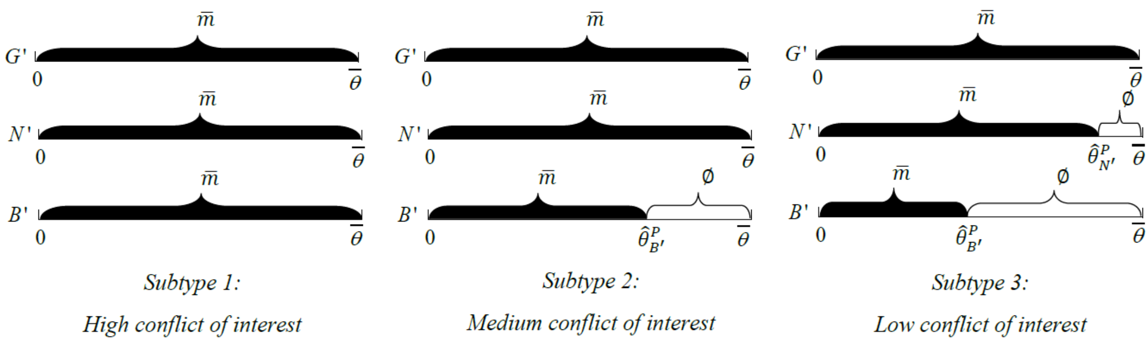

Figure 2: Pooling equilibrium.

whether $\widehat{\theta}_{N'}$ and $\widehat{\theta}_{B'}$ are equal to $\bar{\theta}$, which is in turn determined by the value of F (i.e. the magnitude of the monetary conflict of interest in the bad state of the world).

The pooling equilibrium can be considered as a baseline under outcome-based preferences since one can show that all other existing equilibria are payoff-equivalent. Moreover, there exists no responsive equilibrium where types in states B' and N' induce investment while separating from types in state G'.

Proposition 2 *Under outcome-based preferences:*

(i) All equilibria are payoff-equivalent to the pooling equilibrium.

(ii) There exists no responsive equilibrium where a sender type induces investment in state $i^s \in \{B', N'\}$ with a message not sent by types in state G'.

Part (i) follows from the fact that in all existing equilibria the cutoffs $\widehat{\theta}_{B'}$ and $\widehat{\theta}_{N'}$ are the same (equal to F/ρ and $2F/\rho$, respectively), as noted above. In turn, since under outcome-based preferences the sender's expected utility is fully determined by his type, information state and whether he sends an investment-inducing message or not (which is in turn determined by whether he is below or above the cutoff), the claim follows.

The intuition for part (ii) of the proposition is that a separation of types in states B' and N' from types in state G' would imply that the former types induce beliefs lower than 0.5 with their message. In turn, at least some types in state G' must induce beliefs strictly above 0.5 in any responsive equilibrium (since all of these types send an investment-inducing message by Lemma 3, and hence their relative share among all types inducing investment is the highest). Then, the separating types in states B' and N' would have a strict incentive to deviate to a more persuasive message sent by the types in state G'. Indeed, under outcome-based preferences, such message would also induce investment (and hence yield the same expected utility), while at the same time lead to higher receiver's beliefs (and hence will be strictly preferable in terms of lexicographic preferences by Assumption 2).

3.3 Guilt aversion

The next sections show which equilibria arise if the sender has guilt-averse preferences.

3.3.1 Equilibrium strategies

As noted in Section 2.2.1, the receiver invests after a given sender's action $a^s \in \{M, \varnothing\}$ if and only if $\eta(a^s) > \underline{\eta}$. In turn, the sender chooses his action to maximize his expected utility given the receiver's equilibrium strategy. By (7) and (8), this expected utility is given by

$$U^s_{i^s}(\theta, a^s, a^r) = \begin{cases} 0 & \text{if } \eta(a^s) \leq \underline{\eta} \text{ or } a^s = \varnothing, \\ F - \theta\lambda_{i^s}(E^s E^r[U^r(I)|a^s] - c) & \\ \quad = F - \theta\lambda_{i^s}\eta(a^s)(P - c) & \text{if } \eta(a^s) > \underline{\eta} \text{ and } a^s \in M, \end{cases} \tag{10}$$

where the last equality follows from the consistency of the sender's second-order beliefs in equilibrium (i.e. $E^s E^r[U^r(I)|a^s] = E^r[U^r(I)|a^s] = \eta(a^s)(P - c) + c$). Thus, in case of sending a message, the sender faces a tradeoff between inducing investment by being sufficiently persuasive (to ensure $\eta(m) > \underline{\eta}$), and at the same time keeping the receiver's expectations low to mitigate guilt (since $\eta(m)$ enters negatively in the sender's utility function once the receiver invests).

3.3.2 Pooling equilibrium

As in the case of outcome-based preferences, one can show that there always exists an essentially unique pooling equilibrium also under guilt aversion.

Proposition 3 *There exists an essentially unique pooling equilibrium.*

The basic mechanism behind this equilibrium is as follows: In Subtype 1 of the equilibrium (see Figure 2), the sender's monetary incentive F (i.e. the degree of his conflict of interest) is high enough such that all sender types in all states want to pool on the message $\bar{m}$, which induces investment. If the value of the monetary incentive decreases (Subtypes 2 and 3), then the most guilt-sensitive types in states N' and B' prefer to refrain from advice to avoid guilt. Clearly, the fraction of such types is larger in state B', where the expected guilt is higher than in state N' for a given guilt sensitivity. Besides, no type has a strict incentive to deviate to out-of-equilibrium messages, which lead to to the same receiver's beliefs as $\bar{m}$ does.

The equilibrium beliefs conditional on $\bar{m}$ are determined by the cutoffs $\widehat{\theta}^P_{B'}$ and $\widehat{\theta}^P_{N'}$ (in particular, by substituting ($\Pr[\bar{m}|G'] = 1$, $\Pr[\bar{m}|N'] = \widehat{\theta}^P_{N'}/\bar{\theta}$ and $\Pr[\bar{m}|B'] = \widehat{\theta}^P_{B'}/\bar{\theta}$ into (9)). One can show that lower cutoffs correspond to a higher persuasiveness of the message $\bar{m}$ (i.e. a higher probability of the good state of the world conditional on the message). As Proposition 3 implies, the cutoff types which lead to the persuasiveness of $\bar{m}$ making these types indifferent between $\bar{m}$ and $\varnothing$ (i.e. ensuring the equilibrium), are always unique for given parameter values. Finally, the receiver always finds it optimal to invest following $\bar{m}$ by the same argument as in the case of outcome-based preferences.

Note that in the pooling equilibrium there are two types of loss to the receiver from the ex ante perspective. The first one results from getting message $\bar{m}$ in state B' (which is driven by the monetary bias in sender incentives, termed as *bias-driven damage*).[21] The second type of loss is caused by the sender refraining from advice in state N' if the receiver abstains after this (or *guilt-driven damage*). In this case, since investment is ex ante profitable by Assumption 1, the receiver would strictly prefer to invest instead had she known that the sender is actually uninformed. Such a situation can be interpreted as an inefficient reluctance of the sender to recommend products that are risky though profitable from an ex ante perspective. In terms of the medical example, a doctor who is too afraid of appearing incompetent (or being prosecuted for bad treatment) might prescribe to his patient only the most conservative traditional treatments with predictable but low efficiency, instead on providing advice on more innovative (and hence, more risky), but more promising treatment methods. Analogously, a financial advisor might be reluctant to recommend reasonably risky but profitable financial products.

Finally, note that under a certain range of parameters there also exists a payoff-equivalent equilibrium where the sender types above the cutoff pool on a message which leads to the receiver's abstaining (and hence to a utility of 0 for the sender) instead of refraining from advice. This equilibrium exists whenever the receiver's beliefs induced by such message do not trigger investment (as otherwise the corresponding sender types would rather prefer to avoid guilt by refraining from advice).

3.3.3 Separating equilibrium

Note that generally types in state G' have different preferences over receiver beliefs relative to types in the other states: while the former prefer to send a more persuasive message by lexicographic preferences (while their expected guilt is always zero), sender types in states B' and N' always strictly prefer a less persuasive message. As shown below, such asymmetry gives rise to the possibility of separation between types in state G' on the one side, and types in states B' and N' on the other.

In particular, we define the *separating equilibrium* where, besides the most persuasive message $\bar{m}$ as in the pooling equilibrium, an additional "evasive" message $\widetilde{m}$ is used.

Definition 2 *Separating equilibrium is defined as an equilibrium in which*

- *Some message $\bar{m}$ is sent by all types in state G';*
- *Some other message $\widetilde{m}$ is sent if $\theta \in (0, \widehat{\theta}^S_{B'}]$ in state B', and $\theta \in (0, \widehat{\theta}^S_{N'}]$ in state N';*
- *All other types in states B' and N' (if any) refrain from advice;*
- *The receiver invests after $\bar{m}$ and $\widetilde{m}$, and invests after refrainment if and only if $\eta(\varnothing) > \underline{\eta}$;*

[21] See Sobel (2018) for the introduction of the notion of *damage* in communication games. Our definition is slightly different since the receiver's loss is calculated from the sender's ex ante perspective rather than from the ex post perspective.

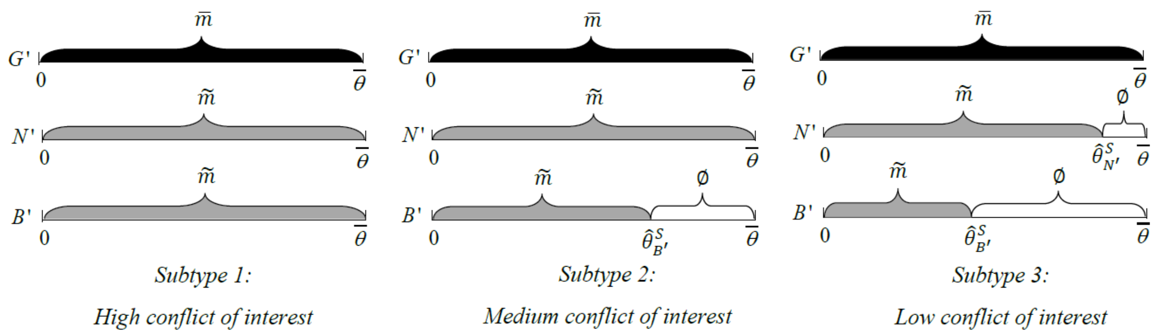

Figure 3: Separating equilibrium.

- *The beliefs after $\bar{m}$, $\widetilde{m}$ and $\varnothing$ are determined by Bayes' rule if possible. For any out-of-equilibrium message $\widehat{m}$ it holds that $\eta(\widehat{m}) \in [\eta(\widetilde{m}), \eta(\bar{m})]$, while the receiver invests after it.*[22]

Let us call a separating equilibrium *essentially unique* if all other existing separating equilibria are the same except for the exogenous value of messages $\bar{m}$ and $\widetilde{m}$. Then, the following holds.

Proposition 4 *There exists an essentially unique separating equilibrium if either of the following holds:*

- $\kappa \in \left(0, \frac{P+c}{P}\right)$*;*
- $\kappa \in \left[\frac{P+c}{P}, \frac{2(P+c)}{2P+c}\right)$ *and* $F < \frac{\bar{\theta}(P+c)(1-\kappa)}{\kappa}$.

Otherwise, there exists no separating equilibrium.

The scheme of this equilibrium is given in Figure 3. Besides types sending the highest message $\bar{m}$ (black figure bracket) and types refraining from advice (white figure bracket) as in the pooling equilibrium, there is a set of types sending the message $\widetilde{m}$ (gray figure bracket).

The intuition for this equilibrium is the following. Sender types in state G', facing no guilt, have the same utility from both $\bar{m}$ and $\widetilde{m}$, which is equal to F since the receiver invests after both messages. At the same time, since $\bar{m}$ leads to higher receiver's beliefs, they strictly prefer it over $\widetilde{m}$ in terms of lexicographic preferences. The types in states N' and B' who induce investment face a strictly positive expected guilt, and hence strictly prefer the evasive message $\widetilde{m}$, since $\eta(\widetilde{m}) < \eta(\bar{m})$ (while the monetary payoff is the same). Thus, the evasive message provides a way to mitigate guilt by inducing less optimistic payoff expectations on the part of the receiver, while still keeping her investing after

[22] As in the case of the pooling equilibrium, (out-of-equilibrium) beliefs conditional on refrainment are not restricted if the latter is not on the equilibrium path.

receiving advice. Besides, no type has an incentive to deviate to any out-of-equilibrium message $\widehat{m}$ as far as the corresponding beliefs are in the specified range.[23]

The receiver's incentive constraints are $\eta(\bar{m}) > \underline{\eta}$ (investment after $\bar{m}$) and $\eta(\widetilde{m}) > \underline{\eta}$ (investment after $\widetilde{m}$). The first constraint is trivially satisfied. The second constraint (meaning that the evasive message is sufficiently credible) is satisfied whenever the share of truly uninformed types is sufficiently high (i.e. κ is sufficiently low):

$$\kappa < \frac{2(P+c)}{2P+c}. \tag{11}$$

In addition, in Subtype 2 of the equilibrium, the only interior cutoff $\widehat{\theta}^S_{B'}$ should be sufficiently distant from the boundary $\bar{\theta}$ (so that there is no excessive pooling of types in state B' pretending to be uninformed), which places additional restriction on the monetary bias F (see the second case in Proposition 4).

Note that the types in state N' who send message $\widetilde{m}$ still experience guilt even though they do not inflate receiver's expectations above their own belief about the state of the world. This can be seen as empirically plausible by the following reasons. First, we abstract away from the exogenous meaning of messages so that $\widetilde{m}$ does not necessarily mean "truth-telling" in its literal sense for the types in state N'. Second, the receiver, after getting message $\widetilde{m}$, can never be sure whether the sender is indeed uninformed, or just pretended to be so while in fact knowing in advance that the investment is unprofitable. Hence, she may still be frustrated with the sender's advice if the investment turns out unsuccessful, which could ultimately cause the corresponding intrinsic or extrinsic costs for the latter.

Regarding inefficiencies arising for the receiver in equilibrium, the separating equilibrium can also feature bias-driven damage (in Subtypes 2 and 3) and guilt-driven damage (in Subtype 3), as in the pooling equilibrium. At the same time, one can show that the cutoffs in the separating equilibrium are higher than in the pooling equilibrium, which generally leads to weakly higher bias-driven damage yet weakly lower guilt-driven damage (see Section 3.3.5 below for the analysis of the corresponding welfare implications).

Lemma 6 *Whenever the separating equilibrium exists, for any $i^s \in \{B', N'\}$ it holds that $\widehat{\theta}^S_{i^s} \geq \widehat{\theta}^P_{i^s}$ with a strict inequality if and only if $\widehat{\theta}^P_{i^s} < \bar{\theta}$.*

The reason for this is that the evasive message $\widetilde{m}$ is strictly less persuasive than the message $\bar{m}$ in the pooling equilibrium, so that a larger fraction of sender types in a given state prefer to induce investment with this message.

3.3.4 Other equilibria

The following proposition justifies the focus of the current analysis on the two types of equilibria considered above.

[23] As in the case of the pooling equilibrium, one can get a payoff-equivalent equilibrium (under a stricter parameter range) where the sender types above the cutoff send a message leading to abstaining instead of refraining from advice.

Proposition 5 *Any existing equilibrium is payoff-equivalent to either the pooling or the separating equilibrium.*

The intuition for this result is that there can be at most two levels of persuasiveness of investment-inducing messages in equilibrium: one for messages sent by types in state G', and one for messages sent by types in states B' and N'. Any heterogeneity within these two groups of messages would imply a strict incentive to deviate for the sender: the types in state G' would always like to deviate to the most persuasive message (according to the lexicographic preferences, see Assumption 2), while the types in the other states would always strictly prefer a less persuasive message (which then yields a lower expected guilt). Hence, either all types inducing investment send messages with the same level of persuasiveness in all states, or the types in state G' separate (inducing $\eta(m) = 1$). In turn, one can show that these two cases would yield the same level of equilibrium persuasiveness of investment-inducing messages as in the pooling and the separating equilibrium, respectively. Finally, the equilibrium level of persuasiveness uniquely pins down the indifferent cutoff types $\widehat{\theta}_{B'}$ and $\widehat{\theta}_{N'}$, which finally leads to payoff equivalence of any existing equilibrium to either the pooling or the separating equilibrium for both players.

Hence, the remaining equilibria are essentially equivalent to either the pooling or the separating equilibrium in the sense of being characterized by the same receiver's beliefs emerging in equilibrium (as far as the corresponding message induces investment). Technically, the only difference of these equilibria from the baseline equilibria considered above is that the former just allow for a larger quantity of differently formulated (investment-inducing) messages on the equilibrium path, which, however, have the same "meaning", i.e. the same receiver's beliefs conditional on the message, as in either the pooling or the separating equilibrium.

3.3.5 Welfare comparison

As shown in Section 3.3.3, guilt aversion on the side of the sender allows for the separating equilibrium, in which the sender types who prefer to conceal their information from the receiver pool with uninformed rather than differently informed types, and which does not arise under purely outcome-based preferences. This section considers whether this structure of equilibrium communication is eventually detrimental or beneficial to the receiver and the sender relative to the pooling equilibrium.

Receiver's utility. Consider the receiver's utility from the ex ante perspective, i.e. before she observes the sender's message. One can show that the receiver's ex ante expected utility in the separating equilibrium can be either higher or lower than in the pooling equilibrium, depending on the monetary conflict of interest.

Proposition 6 *Consider any $\kappa \in \left(0, \frac{2(P+c)}{2P+c}\right)$ so that there exists nonempty interval Φ such that both the pooling and the separating equilibria exist if and only if $F \in \Phi$. Then, there exists $F^* \in \Phi$ such that for any $F \geq F^*$ $(F < F^*)$ where both the pooling and the* *separating equilibria exist, the pooling equilibrium yields a higher (lower) ex ante utility for the receiver than the separating equilibrium, and strictly so for some F.*

The result is based on the fact that the cutoffs in the separating equilibrium are higher (see Lemma 6), since message $\widetilde{m}$ in the separating equilibrium is less persuasive, and hence less costly in terms of guilt than message $\bar{m}$ in the pooling equilibrium. This relation of the cutoffs implies that the rate of bias-driven damage (sending an investment-inducing message in state B') is higher in the separating equilibrium. On the other hand, a lower cutoff in state N' in the pooling equilibrium implies that the rate of guilt-driven damage (refraining from advice in state N', see Figure 2) is higher there. Recall that this is also detrimental to the receiver's utility since she prefers investment over abstaining ex ante (by Assumption 1). Thus, the sender's option to use the evasive message in the separating equilibrium has two effects on the receiver's utility. The negative effect stems from providing psychologically cheap opportunities for the sender to induce investment after observing the bad state by credibly pretending to be uninformed. The positive effect stems from raising the efficiency of communication of uninformed types, whose expected guilt in case of inducing investment is reduced. The total effect depends on which of these two effects dominates.

In particular, if F is sufficiently large, then both pooling and separating equilibria are of either Subtype 1 or Subtype 2, where there is no guilt-driven damage (see Figures 2 and 3). Consequently, the total effect of a switch from the pooling to the separating equilibrium is limited to enhancing bias-driven damage in state B', which leads to a welfare loss for the receiver (except for the case when both equilibria are of Subtype 1 when the receiver's ex ante expected utility does not change). If, to the contrary, F is sufficiently small, both the pooling and the separating equilibrium are of Subtype 3. Then, besides the negative effect, there is an additional positive effect of equilibrium separation (from types in state N') due to a reduction in guilt-driven damage. A clear-cut result here is that the positive effect in this case is always larger than the negative effect related to bias-driven damage.[24] Finally, if F is in the intermediate range where the pooling equilibrium is of Subtype 3 and the separating equilibrium is not, then the comparison between the equilibria depends on how large the effect of guilt-driven damage in the pooling equilibrium is. In particular, there is a unique threshold F^* such that for $F < F^*$ the scope of guilt-driven damage in the pooling equilibrium is large enough to cause an overall lower ex ante utility for the receiver than in the separating equilibrium.

Sender's utility. From the perspective of the sender, the separating equilibrium is

[24] The reason for this is as follows. First, note that a switch from the separating to the pooling equilibrium leads to an overall reduction of investment in states N' and B' (due to the decrease in the cutoffs). At the same time, the expected receiver's payoff *conditional* on obtaining an investment-inducing message in these information states remains the same in both equilibria. This is ensured by the fact that the ratio of the cutoffs in states N' and B' is the same (see Lemmas A.6 and A.9 in the online Appendix). Finally, this conditional expected payoff is positive, because the receiver invests after $\widetilde{m}$ in the separating equilibrium. Hence, the switch from the separating to the pooling equilibrium in this case effectively results (merely) in contraction of ex ante efficient investment, yielding a loss in terms of the ex-ante utility of the receiver.

always the preferable equilibrium.

Proposition 7 *Whenever the separating equilibrium exists, it yields a weakly higher expected utility than the pooling equilibrium for any sender type in any information state, and strictly so at least for some types in states B' and N'.*

Intuitively, any sender types inducing investment in states B' and N' would benefit from the evasive message $\widetilde{m}$ in the separating equilibrium being less persuasive (and hence implying a lower expected guilt) than message $\bar{m}$ in the pooling equilibrium. Moreover, since the cutoffs in the separating equilibrium $\widehat{\theta}^{S}_{i^s}$ are higher, the types between $\widehat{\theta}^{P}_{i^s}$ and $\widehat{\theta}^{S}_{i^s}$ would benefit from switching from refrainment in the pooling equilibrium (which yields a utility of zero) to sending an investment-inducing message in the separating equilibrium (which yields a positive utility). Finally, the types in state G' (always obtaining F) and the types refraining from advice in both equilibria will be indifferent between them.

4 Welfare effects of ex ante information quality

This section analyzes implications of guilt aversion for the players' preferences regarding the sender's ex ante information quality κ (i.e. the ex ante likelihood of him being informed about the state of the world). In particular, we consider whether the receiver is better off when dealing with an ex ante more informed sender, and whether the sender himself ex ante prefers to be better informed. Strikingly, this is not always the case in both instances.

4.1 Receiver's preference over ex ante information quality

Common intuition suggests that once the receiver is rational, and thus cannot be made worse off by communicating with the sender, she should benefit from the sender being more informed. However, as shown below, guilt aversion may cause a negative externality from the sender being more informed for the receiver's welfare under certain conditions. The reason is that for an ex ante less informed (and guilt-averse) sender it is sometimes easier to commit to providing truthful advice.

The following proposition summarizes the comparative statics of the receiver's ex ante expected utility with respect to κ in the pooling and the separating equilibrium (recall that all other existing equilibria are payoff-equivalent to either of them by Proposition 5).

Proposition 8 *(i) In the pooling equilibrium, there exists threshold $\kappa^* \in \left(\frac{4F-\bar{\theta}(P-c)}{F}, 1\right)$ such that the receiver's ex ante utility strictly decreases in κ on some interval K if and only if for any $\kappa \in K$ it holds:*

- $\frac{\bar{\theta}(P-c)}{4} < F < \frac{1}{6}\bar{\theta}(-2c + P + \sqrt{P^2 + 2c(P-c)})$,
- $\kappa \in \left(\frac{4F-\bar{\theta}(P-c)}{F}, \kappa^*\right)$.

(ii) In the separating equilibrium, the receiver's ex ante utility strictly decreases in κ on some interval K if and only if for any $\kappa \in K$ it holds:

- $\max\left\{\frac{\bar{\theta}(P-c)(1-\kappa)}{1+\sqrt{1+\kappa^2}}, \frac{\bar{\theta}(P-c)(1-\kappa)}{4-3\kappa}\right\} < F < \frac{\bar{\theta}(P-c)(1-\kappa)}{2-\kappa}$ *for* $\kappa \in \left(0, \frac{P+c}{P}\right)$,
- $\max\left\{\frac{\bar{\theta}(P-c)(1-\kappa)}{1+\sqrt{1+\kappa^2}}, \frac{\bar{\theta}(P-c)(1-\kappa)}{4-3\kappa}\right\} < F < \frac{\bar{\theta}(P+c)(1-\kappa)}{\kappa}$ *for* $\kappa \in \left[\frac{P+c}{P}, \frac{2(P+c)}{2P+c}\right)$.

Consider the pooling equilibrium. There, an increase in κ has two effects: a direct positive and an indirect negative. The positive effect relates to the fact that once the sender is informed (i.e. is either in state G' or B'), the receiver is (weakly) more likely to invest in the good than in the bad state of the world. The reason for this is that the share of types inducing investment in state B' is (weakly) lower than such share in state G'. On the other hand, if the sender is uninformed, then the receiver is equally likely to invest in both states of the world. Moreover, there can be foregone investment in the good state of the world (due to guilt-driven damage), which never happens with an informed sender. Altogether, this implies that

$$E[U^r|G' \vee B'] \geq E[U^r|N'], \tag{12}$$

so that an increase in κ and, hence, the probability of facing an informed sender has a positive effect on the receiver's welfare (all else equal).

However, an increase in κ has also an indirect negative effect on the receiver's welfare through guilt aversion. In particular, higher κ implies that the message $\bar{m}$ is more likely to be sent by informed types, which by (12) should lead to higher receiver's expectations conditional on the message (all else equal). This results in higher expected guilt of the sender from sending $\bar{m}$, pushing the cutoffs down. In turn, this can cause an increase in guilt-driven damage, which under some parameter values can overweigh the positive effect described in the preceding paragraph.

In the separating equilibrium, an increase in κ again has two effects. The first (positive) effect is the same as in the previous case: all else equal (i.e. for given cutoff values), the receiver prefers to deal with an informed rather than an uninformed sender. The second (negative) effect is driven by guilt aversion: higher κ *decreases* the persuasiveness of the evasive message $\widetilde{m}$. In particular, higher κ implies that the share of truly uninformed types is lower, so that the evasive message $\widetilde{m}$ becomes less credible and rather signals types in state B' who want to conceal their bad news. By being less persuasive, the message $\widetilde{m}$ induces less guilt on the part of the sender so that the equilibrium cutoffs increase. If the equilibrium is of Subtype 2, this leads to a decrease in the receiver's welfare due to the spread of bias-driven damage. Under certain parameter values this can render the total effect of higher κ to be negative.

Thus, the ex ante likelihood of obtaining an informative signal affects the sender's anticipation of guilt, and hence the rate of truth-telling conditional on a given information state. This mechanism leads to seemingly paradoxical cases when even a completely rational receiver prefers to deal with a sender who is less likely to have the information she needs.[25]

4.2 Sender's preference over ex ante information quality

Consider the preference of a given sender type over his likelihood of being informed from the ex ante perspective, i.e. before his information state is realized.[26]

The effect of κ for the sender's ex ante expected utility is summarized in the following proposition.

Proposition 9 *(i) In the pooling equilibrium:*

- *if $F \geq 0.5\bar{\theta}(P - c)$, then the sender is indifferent over his probability of being informed.*
- *if $F < 0.5\bar{\theta}(P-c)$, then the sender strictly prefers ex ante a lower (higher) probability of being informed if his type is sufficiently low (high).*

(ii) In the separating equilibrium, the sender always strictly prefers ex ante a higher probability of being informed.

First, consider the pooling equilibrium. If $F \geq 0.5\bar{\theta}(P - c)$, then the equilibrium is of Subtype 1 (see Figure 2), where all sender types induce investment while the receiver's beliefs conditional on $\bar{m}$ are always equal to her prior of 0.5. In this case, clearly, the sender's ex ante expected utility does not change with κ.

If $F < 0.5\bar{\theta}(P-c)$, then the cutoff $\widehat{\theta}^P_{B'}$ is smaller than $\bar{\theta}$, so that at least some types in state B' should refrain from advice and, hence, condition their decision on the obtained information. Then, for a given level of expected guilt, such types would be strictly better off when knowing the state of the world rather than staying uninformed, in which case they are not able to adjust their behavior to the state. For instance, if some type refrains from advice while being uninformed (and hence gets a utility of 0), he would still induce investment and gain F with probability 0.5 (i.e. in state G') in case if he knows the state.

At the same time, as considered in Section 4.1, the level of κ also affects the receiver's equilibrium beliefs in that higher κ leads to higher beliefs conditional on $\bar{m}$. This is detrimental to the sender's utility by causing higher expected guilt. Whether this negative effects overweighs the positive effect described above depends on the sender's guilt sensitivity. Sufficiently low sender types (in particular, below the cutoff $\widehat{\theta}^P_{B'}$) always induce investment in equilibrium, i.e. do not condition their strategy on the obtained

[25] A positive effect of noise in the sender's information (through a different mechanism than considered here) is also found by Blume et al. (2007) within the benchmark cheap-talk framework of Crawford and Sobel (1982).

[26] Note that a change in κ *after* the sender's information state is realized cannot affect the sender's state of knowledge anymore (and, correspondingly, the receiver's beliefs about it), and thus would not have any effect on the players' welfare.

information, and hence do not benefit from the positive effect of being (objectively) better informed. At the same time, they still suffer from higher guilt stemming from the amended receiver's beliefs conditional on $\bar{m}$ as a result of higher κ. Thus, for these types the overall effect of κ is negative, i.e. they would prefer to commit ex ante to a lower likelihood of being informed. Note that under the range of F and κ specified in Proposition 8(i), *both* the receiver and the lowest sender types would prefer that the ex ante information quality is lower (in the pooling equilibrium).

In contrast, sufficiently high types vary their behavior in equilibrium depending on information state i^s (unless the equilibrium is of Subtype 1). Hence, their benefit from knowing the state of the world is the largest (given also that their utility in state N' tends to be low, or even zero). One can show that for these types this benefit overweighs the loss from higher expected guilt conditional on $\bar{m}$ due to higher κ. Thus, these types would prefer to commit ex ante to the highest possible likelihood of being informed.

Next, consider the separating equilibrium. In the same way as before, the sender benefits from knowing the state of the world and, hence, being able to condition his behavior on this state. Besides, a higher ex ante likelihood that the sender is informed implies that the evasive message $\widetilde{m}$ is less persuasive (since it is then more likely to be sent by types in state B'), in which case the sender has lower expected guilt. Thus, an increase in κ has always a positive effect on the sender's ex ante utility in the separating equilibrium.

4.3 Comparison to the outcome-based model

Notably, the result that the receiver's or the sender's welfare can decline with the ex ante quality of the sender's information κ cannot be explained by outcome-based preferences.

Proposition 10 *Under outcome-based preferences, the ex ante utility of both the receiver and the sender never decreases with κ.*

Intuitively, under outcome-based preferences an increase in κ does not affect the sender's utility from inducing investment in a given information state, and hence the equilibrium cutoffs. The only effect of κ which matters for both players is the resulting change in the distribution of the sender's information states, i.e. the decrease in the relative probability of state N'. As shown in Section 4.1, the receiver then benefits from ex ante more efficient advice, since she is relatively more likely to invest in the good state than in the bad state conditional on the sender being informed. At the same time, as discussed in Section 4.2, each sender type also prefers to be informed (all else equal), since he is then able to adjust his decision to the realized state of the world. Thus, under outcome-based preferences higher κ is beneficial to both players.

5 Discussion

This paper studies a model of strategic communication where the sender is guilt-averse toward the receiver. The main result is that guilt aversion may cause a specific structure of equilibrium communication where the sender types observing the bad state of the world induce investment by pooling with uninformed types, while separating from those types who observe the good state. The emergence of such evasive communication is based on the sender's incentives to reduce the receiver's ex ante payoff expectations as much as possible, while still keeping her choosing the most preferred sender's action (investment). Mimicking uninformed types naturally fits these sender's incentives by being less persuasive than pooling with types observing the good state, yet sufficiently persuasive to induce investment. Notably, such equilibrium is impossible under purely outcome-based preferences if the likelihood of the receiver's investment is even slightly sensitive to her beliefs conditional on the message (which implies at least lexicographic preferences of the sender for higher beliefs of the receiver).

Thus, the model predicts evasive language as a characteristic feature of strategic communication once the sender has guilt-averse preferences.[27] Moreover, it explains under which conditions this feature is detrimental to the receiver (while the sender always benefits from the possibility to send evasive messages in equilibrium). The negative effect of evasion for the receiver stems from providing the sender with a psychologically cheap way to avoid communicating bad news. The positive effect of evasion is driven by the fact that truly uninformed sender types are more prone to induce (ex ante efficient) investment in the separating equilibrium, using the evasive message, than in the pooling equilibrium. Which effect dominates depends on the degree of the monetary conflict of interest, with a high conflict of interest resulting in an overall negative effect of the option to send evasive messages in equilibrium.

Another important consequence of guilt aversion is that higher ex ante quality of the sender's information may backfire for both the receiver and the sender under certain conditions. In particular, the receiver may prefer to deal with an ex ante less informed sender, while the sender himself would sometimes prefer to ex ante commit to a lower quality of his information. The underlying reason for these effects is that the ex ante likelihood of the sender being informed ultimately affects the receiver's beliefs conditional on the equilibrium messages, which in turn changes the level of expected guilt (which is relevant for the sender) and, hence, the sender's incentives to induce investment in different information states (which is relevant for the receiver). Again, such preferences over information cannot be explained by the outcome-based model. Altogether, this suggests that guilt aversion may create adverse incentives with respect to information

[27] Evasive and vague communication is widespread in many social contexts. Examples include academic communication (Metsä-Ketelä 2012), doctor-patient interactions (Adolphs et al. 2007), or corporate annual reports (Guo et al. 2017). See also Serra-Garcia et al. (2011) and Khalmetski et al. (2017) for experimental evidence that senders indeed tend to use evasive and vague messages as a substitute for explicit lying.

transparency in the economy, which opens up a promising direction for future research.[28]

Overall, the results may be applicable to predict patterns of (evasive) communication and the corresponding welfare implications in a range of real-life advice settings where advisors may be driven by guilt aversion toward their customers (e.g., in medical or financial advice). In terms of policy implications, our results suggest that in some situations a policy maker who aims to protect advisees from fraudulent advice may want to restrict advisors' communication to explicit statements (to reduce the scope of evasion and induce a pooling equilibrium structure). At the same time, conventional policies involving punishment of explicit lying (in case it is verifiable ex post) may lead to the emergence of more opaque equilibrium communication (i.e. a separating equilibrium structure), which can eventually backfire for the recipients of advice. Further research may clarify how policy interventions in the sphere of expert advice could be adjusted to possible effects of guilt aversion and, more generally, belief-dependent preferences.

References

Abeler, J., D. Nosenzo, and C. Raymond, forthcoming: "Preferences for Truth-Telling – A Meta Study," Econometrica.

Adolphs, S., S. Atkins, and K. Harvey (2007): "Caught between Professional Requirements and Interpersonal Needs: Vague Language in Healthcare Contexts," In Joan Cutting (ed.), *Vague Language Explored,* 62–78. New York: Palgrave Macmillan.

Anagol, S., S. Cole, and S. Sarkar (2017): "Understanding the Advice of Commissions-Motivated Agents: Evidence from the Indian Life Insurance Market," The Review of Economics and Statistics, 99(1), 1–15.

Austen-Smith, D. (1994): "Strategic Transmission of Costly Information," Econometrica, 62, 955–963.

Battigalli, P., G. Charness, and M. Dufwenberg (2013): "Deception: The Role of Guilt," Journal of Economic Behavior & Organization, 93, 227–232.

Battigalli, P., and M. Dufwenberg (2007): "Guilt in Games," American Economic Review, 97, 170–176.

Battigalli, P., and M. Dufwenberg (2009): "Dynamic Psychological Games," Journal of Economic Theory, 141, 1–35.

Bhattacharya, S., M. Goltsman, and A. Mukherjee (2016): "On the optimality of diverse expert panels in persuasion games," Games and Economic Behavior, 107, 345-363.

[28] Another interesting avenue for future research would be to analyze which forms of evasive communication would arise in a setting with many possible states of the world and a partial conflict of interest, like, e.g., in the model of Crawford and Sobel (1982).

Blume, A., O. J. Board, and K. Kawamura (2007): “Noisy Talk,” Theoretical Economics, 2, 395–440.

Bolton, P., X. Freixas, and J. Shapiro (2007): “Conflicts of Interest, Information Provision, and Competition in Banking,” Journal of Financial Economics, 85, 297–330.

Charness, G., and M. Dufwenberg (2006): “Promises and Partnership,” Econometrica, 74, 1579–1601.

Chater, N., R. Inderst, and S. Huck (2010): “Consumer Decision-Making in Retail Investment Services,” Report to the European Commission.

Crawford, V. P., and J. Sobel (1982): “Strategic Information Transmission,” Econometrica, 50, 1431–1451.

Dufwenberg, M., Jr. and M. Dufwenberg, Sr. (2018): “Lies in Disguise – A Theoretical Analysis of Cheating,” Journal of Economic Theory, 175, 248–264.

Dye, R. A. (1985): “Disclosure of Nonproprietary Information,” Journal of Accounting Research, 23, 123–145.

Dziuda, W. (2011): “Strategic Argumentation,” Journal of Economic Theory, 146, 1362–1397.

Ederer, F., and A. Stremitzer (2017): “Promises and expectations,” Games and Economic Behavior, 106, 161-178.

Ellingsen, T., M. Johannesson, S. Tjøtta, and G. Torsvik (2010): “Testing Guilt Aversion,” Games and Economic Behavior, 68, 95–107.

Erat, S. and U. Gneezy (2012): “White lies,” Management Science, 58(4), 723-733.

Fischbacher, U. and F. Foellmi-Heusi (2013): “Lies in disguise - an experimental study on cheating,” Journal of the European Economic Association, 11(3), 525-547.

Geanakoplos, J., D. Pearce, and E. Stacchetti (1989): “Psychological Games and Sequential Rationality,” Games and Economic Behavior, 1, 60–79.

Gneezy, U. (2005): “Deception: The Role of Consequences,” American Economic Review, 95, 384–394.

Gneezy, U., A. Kajackaite, and J. Sobel (2018): “Lying Aversion and the Size of the Lie,” American Economic Review, 108(2), 419–453.

Guerra, G., and D. J. Zizzo (2004): “Trust Responsiveness and Beliefs,” Journal of Economic Behavior and Organization, 55, 25–30.

Guo, W., T. Yu, and J. Gimeno (2017): “Language and Competition: Communication Vagueness, Interpretation Difficulties, and Market Entry,” Academy of Management Journal, 60(6), 2073–2098.

Inderst, R., and M. Ottaviani (2013): “Sales Talk, Cancellation Terms, and the Role of Consumer Protection,” Review of Economic Studies, 80, 1002–1026.

Johnson, E. M., and M.M. Rehavi (2016): “Physicians Treating Physicians: Information and Incentives in Childbirth,” American Economic Journal: Economic Policy, 8(1), 115-141.

Kartik, N. (2009): “Strategic Communication with Lying Costs,” Review of Economic Studies, 76, 1359–1395.

Kartik, N., M. Ottaviani and F. Squintani (2007): “Credulity, Lies, and Costly Talk,” Journal of Economic Theory, 134, 93–116.

Kesternich, I., H. Schumacher, and J. Winter (2015): “Professional Norms and Physician Behavior: Homo Oeconomicus or Homo Hippocraticus?” Journal of Public Economics, 131, 1–11.

Khalmetski, K. (2016): “Testing Guilt Aversion with an Exogenous Shift in Beliefs,” Games and Economic Behavior, 97, 110–119.

Khalmetski, K., A. Ockenfels, and P. Werner (2015): “Surprising Gifts: Theory and Laboratory Evidence,” Journal of Economic Theory, 159, 163–208.

Khalmetski, K., B. Rockenbach, and P. Werner (2017): “Evasive Lying in Strategic Communication,” Journal of Public Economics, 156, 59–72.

Khalmetski, K., and D. Sliwka, forthcoming: "Disguising Lies - Image Concerns and Partial Lying in Cheating Games,” American Economic Journal: Microeconomics.

Loginova, U. (2012): “Four Essays on Strategic Communication,” Doctoral dissertation, Columbia University.

Lundquist, T., T. Ellingsen, E. Gribbe, and M. Johannesson (2009): “The aversion to lying,” Journal of Economic Behavior & Organization, 70(1-2), 81-92.

McGuire, T. G. (2000): “Physician Agency,” In Handbook of Health Economics, Volume 1A, edited by Anthony J. Culyer and Joseph P. Newhouse, 461–536. Amsterdam; New York and Oxford: Elsevier Science, North-Holland.

Metsä-Ketelä, M. (2012): “Frequencies of Vague Expressions in English as an Academic Lingua Franca,” Journal of English as a Lingua Franca, 1(2), 263–285.

Reuben, E., P. Sapienza, and L. Zingales (2009): “Is Mistrust Self-Fulfilling?” Economics Letters, 104(2), 89–91.

Serra-Garcia, M., Van Damme, E., and J. Potters (2011): "Hiding an Inconvenient Truth: Lies and Vagueness," Games and Economic Behavior, 73(1), 244–261.

Sobel, J. (2018): "Lying and Deception in Games," Working Paper, University of Arizona.

Sutter, M. (2009): "Deception through telling the truth?! Experimental evidence from individuals and teams," The Economic Journal, 119(534), 47-60.

Vanberg, C. (2008): "Why do People Keep their Promises? An Experimental Test of Two Explanations," Econometrica, 76, 1467–1480.

Appendix: Omitted Proofs

A.1 General equilibrium properties

Proof of Lemma 1. Assume by contradiction that there exists an equilibrium where all sender types refrain from advice, and hence get utility of 0. If there exists at least one out-of-equilibrium message inducing receiver's beliefs higher than $\underline{\eta}$ (and hence leading to investment), then at least all types in state G', who have no expected guilt, would deviate to this message obtaining $F > 0$. In the other case, if all out-of-equilibrium messages induce abstaining, then all sender types would strictly prefer any of such messages over refrainment by Assumption 3. ■

Proof of Lemma 2. By Lemma 1 in any equilibrium there should exist at least one message sent with a positive probability. Assume by contradiction that all messages used on the equilibrium path induce conditional beliefs (weakly) lower than $\underline{\eta}$ (so that the receiver abstains after all of them), i.e.

$$\forall m \in \Upsilon_E, \Pr[G|m] \leq \underline{\eta}, \tag{A.1}$$

where Υ_E is the union set of all messages sent with a positive probability in equilibrium. Denote the event that the sender is sending a message $m \in \Upsilon_E$ by $\widetilde{\Upsilon_E}$. Note that since there exists at least one message leading to the receiver's abstaining, no sender type would refrain from advice by Assumption 3 so that

$$\Pr[\widetilde{\Upsilon_E}|G] = \Pr[\widetilde{\Upsilon_E}] = 1.$$

Consequently, by Bayes' rule,

$$\Pr[G|\widetilde{\Upsilon_E}] = \frac{\Pr[\widetilde{\Upsilon_E}|G]\Pr[G]}{\Pr[\widetilde{\Upsilon_E}]} = \Pr[G] = 0.5, \tag{A.2}$$

At the same time, by the law of total probability and the fact that the message space is countable by assumption,

$$\begin{aligned}\Pr[G|\widetilde{\Upsilon_E}] &= \sum_{m\in\Upsilon_E} \Pr[G|m]\Pr[m] \\ &\leq \sum_{m\in\Upsilon_E} \underline{\eta}\Pr[m] = \underline{\eta}\sum_{m\in\Upsilon_E}\Pr[m] = \underline{\eta} = \frac{-c}{P-c} < 0.5, \end{aligned} \tag{A.3}$$

where the first inequality is by (A.1) and the second inequality is by Assumption 1. Since (A.3) contradicts (A.2), the claim follows. ■

Lemma A.1 *A necessary condition for the existence of equilibrium is that the persuasiveness of any investment-inducing message used in states $i^s \in \{B', N'\}$ is the same.*

Proof. Assume by contradiction that there exist two investment-inducing messages m_1 and m_2 sent in at least some of the information states $i^s \in \{B', N'\}$ such that $\eta(m_1) > \eta(m_2)$.

Consider first outcome-based preferences. Since under these preferences the expected utility of the sender is fully determined by his type and the receiver's response, all types inducing investment have the same expected utility conditional on both m_1 and m_2. Consequently, any type θ who sends m_2 in equilibrium (and hence has a positive expected utility if the receiver invests) would like to deviate to m_1 by lexicographic preferences (Assumption 2). Thus, m_2 cannot be an equilibrium message.

Next, consider the case of guilt aversion. If $i^s \neq G'$ (so that $\lambda_{i^s} > 0$) we have

$$U^s_{i^s}(\theta, \eta(m_1), I) = F - \theta\lambda_{i^s}\eta(m_1)(P-c) < F - \theta\lambda_{i^s}\eta(m_2)(P-c) = U^s_{i^s}(\theta, \eta(m_2), I). \quad \text{(A.4)}$$

Then, any type θ would strictly prefer m_2 over m_1 in any information state $i^s \in \{B', N'\}$. Thus, m_1 cannot be an equilibrium message. ■

Proof of Lemma 3. Assume by contradiction that some sender type θ in state G' does not send a message leading to investment so that his expected utility is 0 under both preferences. At the same time, by Lemma 2 there exists at least one message m' leading to investment. By (3) and (10), the expected utility from sending this message in state G' is equal to $F > 0$ under both preferences. Consequently, the sender would have a strict incentive to deviate to m', which yields a contradiction. ■

Proof of Lemma 4. Consider the case of guilt aversion. By Lemma 2 there exists at least one message leading to investment in equilibrium. Next, by Lemma A.1 all messages leading to investment sent by types in states B' and N' have the same persuasiveness, which we denote by η'. Consider further the following possible cases for given $i^s \in \{B', N'\}$.

Case 1: $U^s_{i^s}(\bar{\theta}, I|\eta') > 0$. Then, since $U^s_{i^s}(\theta, I|\eta')$ is continuously decreasing in θ for given η' (see (10)), it follows that for all $\theta < \bar{\theta}$ it holds

$$U^s_{i^s}(\theta, I|\eta') > 0, \quad \text{(A.5)}$$

that is all sender types in state i^s should strictly prefer to send a message leading to investment (which must exist in equilibrium by Lemma 2) over 0 (which would result from sending a message which induces abstaining or from refraining from advice).

Case 2: $U^s_{i^s}(\bar{\theta}, I|\eta') \leq 0$. In this case, since at the same time $U^s_{i^s}(0, I|\eta') = F > 0$, by the intermediate value theorem there must exist type $\widehat{\theta}_{i^s} \in (0, \bar{\theta}]$ such that

$$U^s_{i^s}(\widehat{\theta}_{i^s}, I|\eta') = 0. \quad \text{(A.6)}$$

Then, for all $\theta < \widehat{\theta}_{i^s}$ it holds $U^s_{i^s}(\theta, I|\eta') > 0$. These types strictly prefer to send a message leading to investment over alternative actions. At the same time, for all $\theta > \widehat{\theta}_{i^s}$ (if such types exist, i.e. if $\widehat{\theta}_{i^s} < \bar{\theta}$) it holds $U^s_{i^s}(\theta, I|\eta') < 0$, so that these types strictly prefer to send a message leading to abstaining or to refrain from advice over sending a message leading to investment.

Hence, in all possible cases, if there exists an equilibrium, then it must have the cutoff structure described in the lemma.

The proof for the case of outcome-based preferences follows the same arguments. ■

Proof of Lemma 5. We have

$$\begin{aligned}
\eta(a^s) &\equiv \Pr[G|a^s] = \frac{\Pr[a^s|G]\Pr[G]}{\Pr[a^s]} \\
&= \frac{(\Pr[a^s|G'\cap G]\Pr[G'|G] + \Pr[a^s|N'\cap G]\Pr[N'|G])0.5}{\Pr[a^s|G']\Pr[G'] + \Pr[a^s|N']\Pr[N'] + \Pr[a^s|B']\Pr[B']} \\
&= \frac{(\Pr[a^s|G']\Pr[G'|G] + \Pr[a^s|N']\Pr[N'|G])0.5}{\Pr[a^s|G']\Pr[G'] + \Pr[a^s|N']\Pr[N'] + \Pr[a^s|B']\Pr[B']} \\
&= \frac{\Pr[a^s|G']\kappa + \Pr[a^s|N'](1-\kappa)}{(\Pr[a^s|G'] + \Pr[a^s|B'])\kappa + 2\Pr[a^s|N'](1-\kappa)},
\end{aligned} \tag{A.7}$$

where the second equality is by Bayes' rule, the third equality is by the law of total probability, and the fourth equality is by the fact that equilibrium messages of any sender type θ are fully determined by his information state, i.e. $\Pr[a^s|i^s \cap G] = \Pr[a^s|i^s]$ for $i^s \in \{N', G'\}$. ■

A.2 Outcome-based preferences

Proof of Proposition 1.

Claim 1. *Under outcome-based preferences, the equilibrium cutoffs are given by*

$$\widehat{\theta}_{B'} = \min\{\bar{\theta}, F/\rho\}, \tag{A.8}$$

$$\widehat{\theta}_{N'} = \min\{\bar{\theta}, 2F/\rho\}, \tag{A.9}$$

Proof. Assume by contradiction that this is not the case. Note that by Lemma 4 in any equilibrium the sender induces investment in state $i^s \in \{B', N'\}$ if $\theta < \widehat{\theta}_{i^s}$, and obtains 0 if $\theta > \widehat{\theta}_{i^s}$. Then, some sender types would have a strict incentive to deviate since by (3) the sender prefers the receiver's investment over 0 in state B' (N') if and only if $\theta \le F/\rho$ ($\theta \le 2F/\rho$).

Claim 2. *Under outcome-based preferences, there exists an essentially unique pooling equilibrium.*

Proof. Consider a putative pooling equilibrium with the cutoffs given by (A.8) and (A.9), while the receiver's belief conditional on any out-of-equilibrium message being set at $\eta(\bar{m})$. The utility function (3) implies that the sender prefers a message leading to investment over 0 if and only if he observes the good state or $\theta \le F/\rho$ in state B' or $\theta \le 2F/\rho$ in state N'. Hence, no sender type has an incentive to deviate from sending $\bar{m}$ to refrainment or the other way round given the receiver's response. Besides, no sender type has an incentive to deviate to any out-of-equilibrium message $\widehat{m}$ given that

$\eta(\widehat{m}) = \eta(\bar{m})$ by assumption.

Consider the receiver's incentives to invest after $\bar{m}$ (the receiver's incentive constraint in case of the sender's refrainment is satisfied by construction). By Lemma 5

$$\begin{aligned}\Pr[G|\bar{m}] &= \frac{\Pr[\bar{m}|G']\kappa + \Pr[\bar{m}|N'](1-\kappa)}{(\Pr[\bar{m}|G'] + \Pr[\bar{m}|B'])\kappa + 2\Pr[\bar{m}|N'](1-\kappa)} \\ &\geq \frac{\Pr[\bar{m}|G']\kappa + \Pr[\bar{m}|N'](1-\kappa)}{2\Pr[\bar{m}|G']\kappa + 2\Pr[\bar{m}|N'](1-\kappa)} = 0.5 > \underline{\eta},\end{aligned}$$

where the first inequality follows from the fact that $\Pr[\bar{m}|B'] \leq \Pr[\bar{m}|G']$ by construction of the equilibrium, and the last inequality is by Assumption 1. Hence, the receiver always finds it optimal to invest after $\bar{m}$ (and hence after $\widehat{m}$ given that $\eta(\widehat{m}) = \eta(\bar{m})$).

Thus, all incentive constraints are satisfied so that the prescribed strategies and beliefs constitute an equilibrium.

Finally, by Claim 1, all existing equilibria have the same cutoffs. Hence, all existing pooling equilibria can only be different in terms of exogenous formulation of the messages. Consequently, the initially considered pooling equilibrium is essentially unique. ■

Lemma A.2 *If in equilibrium the sender refrains from advice with a positive probability, then the receiver invests after refrainment if and only if* $\Pr[G|\widetilde{\Theta}_>] > \underline{\eta}$.

Proof. Claim 1. *The sender types strictly above the cutoffs* $\widehat{\theta}_{B'}$ *and* $\widehat{\theta}_{N'}$ *either all refrain from advice or all send message(s) leading to abstaining.*

Proof. Assume by contradiction that some sender types above the cutoffs send message(s) leading to abstaining (which leads to the sender's utility of 0 under both preferences), and some refrain from advice (which also leads to 0). Then, the latter types would have a strict incentive to deviate to any message leading to abstaining by lexicographic preferences (Assumption 3).

Claim 2. If $\Pr[G|\widetilde{\Theta}_>] \leq \underline{\eta}$ *, then the receiver abstains in equilibrium in case if* $\theta > \widehat{\theta}_{i^s}$.

Proof. By Claim 1, the sender types above the cutoff either always send a message leading to abstaining, or all refrain from advice. In the first case, the claim follows immediately. In the second case, the receiver's belief conditional on refraining is

$$\eta(\varnothing) = \Pr[G|\varnothing] = \Pr[G|\widetilde{\Theta}_>] \leq \underline{\eta},$$

where the inequality is by initial assumption. Consequently, the receiver abstains as well.

Claim 3. *If* $\Pr[G|\widetilde{\Theta}_>] > \underline{\eta}$*, then the sender types strictly above the cutoffs always refrain from advice in equilibrium.*

Proof. Assume by contradiction that $\Pr[G|\widetilde{\Theta}_>] > \underline{\eta}$ and there exists type $\theta > \widehat{\theta}_{i^s}$ who sends a message leading to abstaining in equilibrium (messages leading to investment cannot be sent by types $\theta > \widehat{\theta}_{i^s}$ in equilibrium by Lemma 4). Then, by Claim 1, all sender types above the cutoff send a message leading to abstaining. Denote the set of all messages sent by types above the cutoffs in states B' and N' by $\Upsilon_>$. Denote the event that the sender's message belongs to $\Upsilon_>$ as $\widetilde{\Upsilon}_>$. Note that by Lemma 4 any message from $\Upsilon_>$ is never sent by the types below the cutoff, which implies

$$\Pr[G|\widetilde{\Theta}_>] = \Pr[G|\widetilde{\Upsilon}_>].$$

Consequently, we have

$$\begin{aligned}
\Pr[G|\widetilde{\Theta}_>] &= \Pr[G|\widetilde{\Upsilon}_>] = \sum_{m\in\Upsilon_>} \Pr[G|m\cap\widetilde{\Upsilon}_>]\Pr[m|\widetilde{\Upsilon}_>] \\
&= \sum_{m\in\Upsilon_>} \Pr[G|m]\Pr[m|\widetilde{\Upsilon}_>] \leq \sum_{m\in\Upsilon_>} \underline{\eta}\Pr[m|\widetilde{\Upsilon}_>] = \underline{\eta},
\end{aligned}$$

where the second equality is by the law of total probability and the fact that the overall message set is countable by assumption, and the inequality is due to the fact that all messages in $\Upsilon_>$ must induce abstaining. In turn, this contradicts the initial assumption of $\Pr[G|\widetilde{\Theta}_>] > \underline{\eta}$.

Claim 4. *If* $\Pr[G|\widetilde{\Theta}_>] > \underline{\eta}$, *then the receiver invests after* $\varnothing$ *if on the equilibrium path.*

Proof. If $\Pr[G|\widetilde{\Theta}_>] > \underline{\eta}$, then by Claim 3 the sender types above the cutoff must all refrain in equilibrium (while the types below the cutoff never refrain by Lemma 4). Consequently,

$$\Pr[G|\varnothing] = \Pr[G|\widetilde{\Theta}_>] > \underline{\eta},$$

where the inequality is by assumption. Hence, the receiver would always prefer to invest conditional on the sender refraining from advice.

The claim of the lemma follows jointly from Claims 2 and 4. ■

Proof of Proposition 2.

Claim 1. *All equilibria are payoff-equivalent to an existing pooling equilibrium.*

Proof. Consider any existing equilibrium. Let us show that it is payoff-equivalent to a pooling equilibrium (existing by Proposition 1).

By Claim 1 in the previous proof the cutoffs $\widehat{\theta}^P_{B'}$ and $\widehat{\theta}^P_{N'}$ in any two equilibria are the same. Hence, by Lemmas 4 and A.2, the receiver's action conditional on a given sender's type in a given information state (except for the zero measured cutoff types) must be the same between the equilibria. Consequently, the receiver's expected payoff is also the same.

Consider any given sender's type θ in an information state i^s. Note first that the cutoffs $\widehat{\theta}^P_{i^s}$ for $i^s \in \{B', N'\}$ are the same between the equilibria by Claim 1 in the previous proof. Then, we can have the following cases:

- If $i^s \in \{B', N'\}$ while $\theta > \widehat{\theta}^P_{i^s}$, then by Lemma 4 the sender obtains 0 in both equilibria.

- If $i^s \in \{B', N'\}$ while $\theta < \widehat{\theta}^P_{i^s}$, then by Lemma 4 the sender obtains the same expected payoff of $F - \Pr[B|i^s] \cdot \theta\rho$ in both equilibria.

- If $i^s \in \{B', N'\}$ while $\theta = \widehat{\theta}^P_{i^s} < \bar{\theta}$, then the sender must be indifferent between 0 and sending an investment-inducing message, and thus expects 0 in either equilibrium.

- If $i^s \in \{B', N'\}$ while $\theta = \widehat{\theta}^P_{i^s} = \bar{\theta}$, then the sender is either again indifferent between 0 and sending an investment-inducing message, or has a strictly positive utility from inducing investment in both equilibria obtaining the same expected payoff of $F - \Pr[B|i^s] \cdot \theta\rho$.

- If $i^s = G'$, then by Lemma 3 the sender always obtains F in both equilibria.

Hence, any given sender's type in any information state has the same expected utility in both equilibria.

Claim 2. *In any given equilibrium all messages inducing investment have the same persuasiveness.*

Proof. Assume by contradiction that there are two distinct messages m' and m'' on the equilibrium path which induce investment while $\eta(m') > \eta(m'')$. The utility of any given type from sending these messages given by (3) is the same. Then, all types sending m'' in equilibrium (whose expected utility conditional on investment must be positive, since otherwise they would deviate to refraining from advice) would strictly prefer to deviate to m' by the lexicographic preferences (Assumption 2).

Claim 3. *There exists no responsive equilibrium where a sender type induces investment in state $i^s \in \{B', N'\}$ with a message not sent by types in state G' .*

Proof. Assume by contradiction that this is the case. Note that all message inducing investment must lead to the same persuasiveness η' by Claim 2. Since some messages induce investment while being sent only by types in states B' and/or N' by assumption (while $\Pr[G|B'] = 0 < \Pr[G|N'] = 0.5$), we must have

$$\eta' \leq 0.5. \tag{A.10}$$

Note also that since the equilibrium is assumed to be responsive (i.e. the ex ante likelihood of investment is smaller than 1), then by Lemma 4 and the fact that $\widehat{\theta}_{B'} \leq \widehat{\theta}_{N'}$ (see (A.8) and (A.9)) we have

$$\widehat{\theta}_{B'} < \bar{\theta}. \tag{A.11}$$

Next, denote by Υ_I the set of all messages inducing investment in equilibrium. Denote by $\widetilde{\Upsilon_I}$ the event that the sender's message belongs to Υ_I. By the law of total probability and the fact that the message space is assumed to be countable, we have

$$\begin{aligned}\Pr[G|\widetilde{\Upsilon_I}] &= \sum_{m\in\Upsilon_I} \Pr[G|m \cap \widetilde{\Upsilon_I}] \Pr[m|\widetilde{\Upsilon_I}] \\ &= \sum_{m\in\Upsilon_I} \Pr[G|m] \Pr[m|\widetilde{\Upsilon_I}] = \sum_{m\in\Upsilon_I} \eta' \Pr[m|\widetilde{\Upsilon_I}] = \eta'.\end{aligned} \tag{A.12}$$

At the same time, it holds

$$\begin{aligned}\Pr[G|\widetilde{\Upsilon}_I] &= \frac{\Pr[\widetilde{\Upsilon}_I|G']\kappa + \Pr[\widetilde{\Upsilon}_I|N'](1-\kappa)}{(\Pr[\widetilde{\Upsilon}_I|G'] + \Pr[\widetilde{\Upsilon}_I|B'])\kappa + 2\Pr[\widetilde{\Upsilon}_I|N'](1-\kappa)} \\ &> \frac{\Pr[\widetilde{\Upsilon}_I|G']\kappa + \Pr[\widetilde{\Upsilon}_I|N'](1-\kappa)}{2\Pr[\widetilde{\Upsilon}_I|G']\kappa + 2\Pr[\widetilde{\Upsilon}_I|N'](1-\kappa)} = 0.5, \end{aligned} \tag{A.13}$$

where the first equality follows by Bayes' rule by the same derivations as in (A.7), and the inequality is due to $\Pr[\widetilde{\Upsilon}_I|B'] < \Pr[\widetilde{\Upsilon}_I|G']$ by Lemma 3 and (A.11). (A.12) and (A.13) finally lead to

$$\eta' > 0.5,$$

which contradicts (A.10).

The statement of the proposition follows from Claims 1 and 3. ■

A.3 Guilt aversion: General equilibrium properties

Lemma A.3 *A necessary condition for the existence of equilibrium is* $\widehat{\theta}_{N'} \geq \widehat{\theta}_{B'}$.

Proof. Assume by contradiction that there exists an equilibrium such that

$$\widehat{\theta}_{N'} < \widehat{\theta}_{B'}. \tag{A.14}$$

Then, by Lemma 4 all types in the interval $(\widehat{\theta}_{N'}, \widehat{\theta}_{B'})$ obtain 0 in state N' and induce investment in state B'. Consequently, these types have a negative utility from any investment-inducing message in state N' (otherwise they would deviate to such message), and a positive utility from inducing investment in state B' (otherwise, they would deviate to $\varnothing$). Hence, for any $\theta' \in (\widehat{\theta}_{N'}, \widehat{\theta}_{B'})$ it holds

$$U^s_{N'}(\theta', \widehat{\eta}, I) < 0 \leq U^s_{B'}(\theta', \widehat{\eta}, I), \tag{A.15}$$

where $\widehat{\eta}$ is the persuasiveness of investment-inducing messages in states $i^s \in \{B', N'\}$ (unique by Lemma A.1). At the same time, for any $\theta \in (0, \bar{\theta}]$ it holds (given (10))

$$U^s_{B'}(\theta, \widehat{\eta}, I) = F - \theta\widehat{\eta}(P - c) < F - 0.5\theta\widehat{\eta}(P - c) = U^s_{N'}(\theta, \widehat{\eta}, I), \tag{A.16}$$

which yields a contradiction to (A.15). ■

Lemma A.4 *A necessary condition for the existence of equilibrium is that either* $\widehat{\theta}_{i^s} < \bar{\theta}$ *and* $U^s_{i^s}(\widehat{\theta}_{i^s}, \eta(m(\widehat{\theta}_{i^s})), I) = 0$ *or* $\widehat{\theta}_{i^s} = \bar{\theta}$ *and* $U^s_{i^s}(\bar{\theta}, \eta(m(\widehat{\theta}_{i^s})), I) \geq 0$.

Proof. Denote by $\widehat{\eta}$ the persuasiveness of investment-inducing messages in states $i^s \in \{B', N'\}$ (unique by Lemma A.1). Assume by contradiction that equilibrium exists and either $\widehat{\theta}_{i^s} < \bar{\theta}$ and $U^s_{i^s}(\widehat{\theta}_{i^s}, \widehat{\eta}, I) \neq 0$ or $\widehat{\theta}_{i^s} = \bar{\theta}$ and $U^s_{i^s}(\widehat{\theta}_{i^s}, \widehat{\eta}, I) < 0$ for some $i^s \in \{B', N'\}$. Let us demonstrate a contradiction in each of these cases separately.

Case 1: $\widehat{\theta}_{i^s} < \bar{\theta}$ and $U^s_{i^s}(\widehat{\theta}_{i^s}, \widehat{\eta}, I) \neq 0$.

First, let us consider the case $U^s_{i^s}(\widehat{\theta}_{i^s}, \widehat{\eta}, I) > 0$. Then, since $U^s_{i^s}(\theta, \widehat{\eta}, I)$ is continuously decreasing in θ (see (10)), there would exist types sufficiently close to $\widehat{\theta}_{i^s}$ such that $\theta > \widehat{\theta}_{i^s}$ and $U^s_{i^s}(\theta, \widehat{\eta}, I) > 0$. At the same time, according to Lemma 4 such types get the utility of 0 in equilibrium, so that they would have a strict incentive to deviate to an investment-inducing message yielding the positive utility $U^s_{i^s}(\theta, \widehat{\eta}, I)$. Analogously, if $U^s_{i^s}(\widehat{\theta}_{i^s}, \widehat{\eta}, I) < 0$, then types sufficiently close to $\widehat{\theta}_{i^s}$ on the left have a negative utility from inducing investment and hence a strict incentive to deviate to refraining from advice. We have come to contradiction in all possible cases.

Case 2: $\widehat{\theta}_{i^s} = \bar{\theta}$ and $U^s_{i^s}(\bar{\theta}, \widehat{\eta}, I) < 0$.

Then, analogously to the previous case, there would exist types θ such that $\theta < \widehat{\theta}_{i^s}$ and $U^s_{i^s}(\theta, \widehat{\eta}, I) < 0$, which then would like to deviate from sending an investment-inducing message to refraining that yields a contradiction. ■

Lemma A.5 *A necessary condition for the existence of equilibrium is* $\widehat{\theta}_{N'} = \min\left\{\bar{\theta}, 2\widehat{\theta}_{B'}\right\}$.

Proof. Denote by $\widehat{\eta}$ the persuasiveness of investment-inducing messages in states $i^s \in \{B', N'\}$ (unique by Lemma A.1). Assume by contradiction that an equilibrium exists and $\widehat{\theta}_{N'} \neq \min\left\{\bar{\theta}, 2\widehat{\theta}_{B'}\right\}$. Let us consider two possible cases $\widehat{\theta}_{N'} < \bar{\theta}$ and $\widehat{\theta}_{N'} = \bar{\theta}$.

Case 1: $\widehat{\theta}_{N'} < \bar{\theta}$. By Lemma A.3 it then follows $\widehat{\theta}_{B'} < \bar{\theta}$. Consequently, by Lemma A.4 and equation (10) we have

$$U^s_{B'}(\widehat{\theta}_{B'}, \widehat{\eta}, I) = F - \widehat{\theta}_{B'}\widehat{\eta}(P - c) = 0, \tag{A.17}$$
$$U^s_{N'}(\widehat{\theta}_{N'}, \widehat{\eta}, I) = F - 0.5\widehat{\theta}_{N'}\widehat{\eta}(P - c) = 0. \tag{A.18}$$

This implies

$$\widehat{\theta}_{B'} = \frac{F}{\widehat{\eta}(P - c)}, \tag{A.19}$$
$$\widehat{\theta}_{N'} = \frac{F}{0.5\widehat{\eta}(P - c)}, \tag{A.20}$$

and, consequently,

$$\widehat{\theta}_{N'} = 2\widehat{\theta}_{B'}. \tag{A.21}$$

This, together with $\widehat{\theta}_{N'} < \bar{\theta}$, implies $\min\left\{\bar{\theta}, 2\widehat{\theta}_{B'}\right\} = 2\widehat{\theta}_{B'}$, which finally yields a contradiction to $\widehat{\theta}_{N'} \neq \min\left\{\bar{\theta}, 2\widehat{\theta}_{B'}\right\}$.

Case 2: $\widehat{\theta}_{N'} = \bar{\theta}$. Since we have also assumed $\widehat{\theta}_{N'} \neq \min\left\{\bar{\theta}, 2\widehat{\theta}_{B'}\right\}$, this implies

$$\bar{\theta} \neq \min\left\{\bar{\theta}, 2\widehat{\theta}_{B'}\right\} \tag{A.22}$$
$$\Rightarrow \widehat{\theta}_{B'} < 0.5\bar{\theta}. \tag{A.23}$$

Next, we have

$$\begin{aligned} U^s_{N'}(\bar{\theta},\widehat{\eta},I) &= F - 0.5\bar{\theta}\widehat{\eta}(P-c) \\ &= U^s_{B'}(0.5\bar{\theta},\widehat{\eta},I) < 0, \end{aligned} \tag{A.24}$$

where the inequality holds due to $\widehat{\theta}_{B'} < 0.5\bar{\theta}$ by (A.23) and the fact that $U^s_{i^s}(\theta,\widehat{\eta},I)$ is strictly decreasing in θ. Consequently, types sufficiently close to $\bar{\theta}$ in state N' would strictly prefer refrainment over an investment-inducing message which contradicts $\widehat{\theta}_{N'} = \bar{\theta}$.
■

A.4 Pooling equilibrium

Lemma A.6 *If the strategies and beliefs are specified according to Definition 1, then the sender has no incentives to deviate if and only if the following two conditions hold:*
1) either $\widehat{\theta}^P_{B'} < \bar{\theta}$ and $U^s_{B'}(\widehat{\theta}^P_{B'},\eta(\bar{m}),I) = 0$ or $\widehat{\theta}^P_{B'} = \bar{\theta}$ and $U^s_{B'}(\bar{\theta},\eta(\bar{m}),I) \geq 0$.
2) $\widehat{\theta}^P_{N'} = \min\left\{\bar{\theta}, 2\widehat{\theta}^P_{B'}\right\}$.

Proof. The necessity of both conditions follows by the arguments given in the proofs of Lemmas A.4 and A.5. Let us consider their sufficiency and assume that both conditions hold. First, no type in state G' has an incentive to deviate to refrainment while his utility from investment F is strictly positive. Second, it is straightforward to show that, once the first condition holds, no type in state B' has a strict incentive to deviate to refrainment given that $U^s_{B'}(\widehat{\theta}^P_{B'},\eta(\bar{m}),I)$ is continuously decreasing in its first argument.

Let us show that all types in state N' also do not have incentives to deviate to refrainment. Consider the following possible cases.

Case 1: $\widehat{\theta}^P_{B'} < 0.5\bar{\theta}$. Then, by assumption, $\widehat{\theta}^P_{N'} = 2\widehat{\theta}^P_{B'} < \bar{\theta}$. Hence,

$$\begin{aligned} U^s_{N'}(\widehat{\theta}^P_{N'},\eta(\bar{m}),I) &= U^s_{N'}(2\widehat{\theta}^P_{B'},\eta(\bar{m}),I) \\ &= F - \widehat{\theta}^P_{B'}\eta(\bar{m})(P-c) \\ &= U^s_{B'}(\widehat{\theta}^P_{B'},\eta(\bar{m}),I) = 0, \end{aligned} \tag{A.25}$$

where the last equality follows from $\widehat{\theta}^P_{B'} < 0.5\bar{\theta}$ and the first condition of the lemma. Then, by the same arguments as in the case of state B', no sender types in state N' have an incentive to deviate to 0.

Case 2: $\widehat{\theta}^P_{B'} \geq 0.5\bar{\theta}$. Then, by assumption, $\widehat{\theta}^P_{N'} = \bar{\theta}$. This yields

$$\begin{aligned} U^s_{N'}(\widehat{\theta}^P_{N'},\eta(\bar{m}),I) &= U^s_{N'}(\bar{\theta},\eta(\bar{m}),I) \\ &= F - 0.5\bar{\theta}\eta(\bar{m})(P-c) \\ &\geq F - \widehat{\theta}^P_{B'}\eta(\bar{m})(P-c) \\ &= U^s_{B'}(\widehat{\theta}^P_{B'},\eta(\bar{m}),I) \geq 0, \end{aligned} \tag{A.26}$$

where the first inequality follows from $\widehat{\theta}^{P}_{B'} \geq 0.5\bar{\theta}$ and the second one from the first condition of the lemma. Then, for all $\theta \leq \widehat{\theta}^{P}_{N'}$ it holds $U^{s}_{N'}(\theta, \eta(\bar{m}), I) \geq 0$, so that no sender type in state N' has an incentive to deviate to 0.

Finally, no type in either state has a strict incentive to deviate to out-of-equilibrium messages which are equally persuasive as $\bar{m}$ (see Definition 1). ■

Corollary A.1 *In any pooling equilibrium,* $\widehat{\theta}^{P}_{B'} = \min\left\{\frac{F}{\eta(\bar{m})(P-c)}, \bar{\theta}\right\}$.

Proof. Note that $\min\left\{\frac{F}{\eta(\bar{m})(P-c)}, \bar{\theta}\right\} = \bar{\theta}$ if and only if $U^{s}_{B'}(\bar{\theta}, \eta(\bar{m}), I) \geq 0$. Then, the result follows from Lemma A.6. ■

Lemma A.7 *If the strategies and beliefs are specified as in Definition 1 and* $\widehat{\theta}^{P}_{N'} = \min\left\{\bar{\theta}, 2\widehat{\theta}^{P}_{B'}\right\}$*, then* $U^{s}_{B'}(\widehat{\theta}^{P}_{B'}, \eta(\bar{m}|\widehat{\theta}^{P}_{B'}), I)$ *is continuous and strictly decreasing in* $\widehat{\theta}^{P}_{B'}$ *on* $(0, \bar{\theta}]$.

Proof. The continuity of $U^{s}_{B'}(\widehat{\theta}^{P}_{B'}, \eta(\bar{m}|\widehat{\theta}^{P}_{B'}), I)$ follows from the continuity of $\eta(\bar{m}|\widehat{\theta}^{P}_{B'})$ in $\widehat{\theta}^{P}_{B'}$, given also that $\widehat{\theta}^{P}_{N'} = \min\left\{\bar{\theta}, 2\widehat{\theta}^{P}_{B'}\right\}$ is continuous in $\widehat{\theta}^{P}_{B'}$. Note that $\eta(\bar{m}|\widehat{\theta}^{P}_{B'})$ and hence $U^{s}_{B'}(\widehat{\theta}^{P}_{B'}, \eta(\bar{m}|\widehat{\theta}^{P}_{B'}), I)$ are differentiable at all $\widehat{\theta}^{P}_{B'} \in (0, \bar{\theta}]$ except for $\widehat{\theta}^{P}_{B'} = \bar{\theta}$ (the endpoint) and $\widehat{\theta}^{P}_{B'} = 0.5\bar{\theta}$ (since $\widehat{\theta}^{P}_{N'} = \min\left\{\bar{\theta}, 2\widehat{\theta}^{P}_{B'}\right\}$ is not differentiable at this point). For all other values of $\widehat{\theta}^{P}_{B'}$, it holds

$$\frac{dU^{s}_{B'}(\widehat{\theta}^{P}_{B'}, \eta(\bar{m}|\widehat{\theta}^{P}_{B'}), I)}{d\widehat{\theta}^{P}_{B'}} = \frac{\partial U^{s}_{B'}}{\partial \widehat{\theta}^{P}_{B'}} + \frac{\partial U^{s}_{B'}}{\partial \eta}\frac{\partial \eta(\bar{m}|\widehat{\theta}^{P}_{B'})}{\partial \widehat{\theta}^{P}_{B'}}. \tag{A.27}$$

Substituting for $U^{s}_{B'}$ from (10) we get

$$\frac{dU^{s}_{B'}(\widehat{\theta}^{P}_{B'}, \eta(\bar{m}|\widehat{\theta}^{P}_{B'}), I)}{d\widehat{\theta}^{P}_{B'}} = -(P-c)\widehat{\theta}^{P}_{B'}\left(\frac{\partial \eta(\bar{m}|\widehat{\theta}^{P}_{B'})}{\partial \widehat{\theta}^{P}_{B'}} + \frac{\eta(\bar{m}|\widehat{\theta}^{P}_{B'})}{\widehat{\theta}^{P}_{B'}}\right). \tag{A.28}$$

Further, consider the following possible cases given that $\widehat{\theta}^{P}_{N'} = \min\left\{\bar{\theta}, 2\widehat{\theta}^{P}_{B'}\right\}$.

Case 1: $\widehat{\theta}^{P}_{B'} \in (0, 0.5\bar{\theta})$, $\widehat{\theta}^{P}_{N'} = 2\widehat{\theta}^{P}_{B'}$.
Then, by (9) (substituting $\Pr[\bar{m}|G'] = 1$, $\Pr[\bar{m}|N'] = 2\widehat{\theta}^{P}_{B'}/\bar{\theta}$ and $\Pr[\bar{m}|B'] = \widehat{\theta}^{P}_{B'}/\bar{\theta}$)

$$\eta(\bar{m}|\widehat{\theta}^{P}_{B'}) = \frac{2\widehat{\theta}^{P}_{B'}(1-\kappa) + \kappa\bar{\theta}}{\widehat{\theta}^{P}_{B'}(4-3\kappa) + \kappa\bar{\theta}}. \tag{A.29}$$

This function is convex in $\widehat{\theta}^{P}_{B'}$:

$$\frac{\partial^2 \eta(\bar{m}|\widehat{\theta}^{P}_{B'})}{\partial (\widehat{\theta}^{P}_{B'})^2} = \frac{2(2-\kappa)(4-3\kappa)\kappa\bar{\theta}}{((4-3\kappa)\widehat{\theta}^{P}_{B'} + \kappa\bar{\theta})^3} > 0. \tag{A.30}$$

The convexity implies

$$\frac{\partial\eta(\bar{m}|\widehat{\theta}^P_{B'})}{\partial\widehat{\theta}^P_{B'}} > \frac{\eta(\bar{m}|\widehat{\theta}^P_{B'}) - \eta(\bar{m}|0)}{\widehat{\theta}^P_{B'}} = \frac{\eta(\bar{m}|\widehat{\theta}^P_{B'}) - 1}{\widehat{\theta}^P_{B'}}, \tag{A.31}$$

where the last equality is by (A.29). Then, coming back to (A.28) we have

$$\begin{aligned} & \frac{\partial\eta(\bar{m}|\widehat{\theta}^P_{B'})}{\partial\widehat{\theta}^P_{B'}} + \frac{\eta(\bar{m}|\widehat{\theta}^P_{B'})}{\widehat{\theta}^P_{B'}} \\ > \; & \frac{\eta(\bar{m}|\widehat{\theta}^P_{B'}) - 1}{\widehat{\theta}^P_{B'}} + \frac{\eta(\bar{m}|\widehat{\theta}^P_{B'})}{\widehat{\theta}^P_{B'}} \\ = \; & \frac{2\eta(\bar{m}|\widehat{\theta}^P_{B'}) - 1}{\widehat{\theta}^P_{B'}} = \frac{1}{\widehat{\theta}^P_{B'}} \frac{\kappa(\bar{\theta} - \widehat{\theta}^P_{B'})}{\widehat{\theta}^P_{B'}(4-3\kappa) + \bar{\theta}\kappa} > 0, \end{aligned} \tag{A.32}$$

where the first inequality is by (A.31) and the last equality by (A.29). Taken together, (A.32) and (A.28) lead to the claim for $\widehat{\theta}^P_{B'} \in (0, 0.5\bar{\theta})$.

Case 2: $\widehat{\theta}^P_{B'} \in (0.5\bar{\theta}, \bar{\theta})$, $\widehat{\theta}^P_{N'} = \bar{\theta}$.

In this case, $\Pr[\bar{m}|G'] = 1$, $\Pr[\bar{m}|N'] = 1$ and $\Pr[\bar{m}|B'] = \widehat{\theta}^P_{B'}/\bar{\theta}$ so that (9) implies

$$\eta(\bar{m}|\widehat{\theta}^P_{B'}) = \frac{\bar{\theta}}{\kappa\widehat{\theta}^P_{B'} + (2-\kappa)\bar{\theta}}. \tag{A.33}$$

Then, it is possible to obtain a simple closed-form solution for the RHS of (A.28):

$$\begin{aligned} & \frac{\partial\eta(\bar{m}|\widehat{\theta}^P_{B'})}{\partial\widehat{\theta}^P_{B'}} + \frac{\eta(\bar{m}|\widehat{\theta}^P_{B'})}{\widehat{\theta}^P_{B'}} \\ = \; & -\frac{\kappa\bar{\theta}}{(\kappa\widehat{\theta}^P_{B'} + (2-\kappa)\bar{\theta})^2} + \frac{\bar{\theta}}{\widehat{\theta}^P_{B'}(\kappa\widehat{\theta}^P_{B'} + (2-\kappa)\bar{\theta})} \\ = \; & \frac{(2-\kappa)\bar{\theta}^2}{\widehat{\theta}^P_{B'}(\kappa\widehat{\theta}^P_{B'} + (2-\kappa)\bar{\theta})^2} > 0. \end{aligned} \tag{A.34}$$

Taken together, (A.34) and (A.28) lead to the claim for for $\widehat{\theta}^P_{B'} \in (0.5\bar{\theta}, \bar{\theta})$.

Finally, since $U^s_{B'}(\widehat{\theta}^P_{B'}, \eta(\bar{m}|\widehat{\theta}^P_{B'}), I)$ is continuous at all points as noted in the beginning, the claim holds for the whole interval $(0, \bar{\theta}]$. ■

Lemma A.8 *If the strategies and beliefs are specified as in Definition 1, then for any given parameter values there always exist unique cutoffs* $\widehat{\theta}^P_{B'}$ *and* $\widehat{\theta}^P_{N'}$ *such that the sender does not have incentives to deviate. Moreover:*

1) If $F \geq 0.5\bar{\theta}(P-c)$ *then* $\widehat{\theta}^P_{B'} = \widehat{\theta}^P_{N'} = \bar{\theta}$.

2) If $F \in [\bar{\theta}\frac{P-c}{4-\kappa}, 0.5\bar{\theta}(P-c))$ *then* $\widehat{\theta}^P_{B'} \in [0.5\bar{\theta}, \bar{\theta})$ *and* $\widehat{\theta}^P_{N'} = \bar{\theta}$.

3) If $F < \bar{\theta}\frac{P-c}{4-\kappa}$ *then* $\widehat{\theta}^P_{B'} \in (0, 0.5\bar{\theta})$ *and* $\widehat{\theta}^P_{N'} = 2\widehat{\theta}^P_{B'}$.

Proof. For notational simplicity denote

$$\varpi(\widehat{\theta}^P_{B'}) = U^s_{B'}(\widehat{\theta}^P_{B'}, \eta(\bar{m}|\widehat{\theta}^P_{B'}), I) \tag{A.35}$$

with $\widehat{\theta}^P_{N'} = \min\left\{\bar{\theta}, 2\widehat{\theta}^P_{B'}\right\}$. Consider the cases listed in the lemma.

Case 1: $F \geq 0.5\bar{\theta}(P-c)$. Consider the properties of $\varpi(\cdot)$ at the exterior point $\bar{\theta}$. First, note that

$$\eta(\bar{m}|\widehat{\theta}^P_{B'} = \bar{\theta}) = 0.5, \tag{A.36}$$

which results from substituting $\Pr[\bar{m}|G'] = \Pr[\bar{m}|N'] = \Pr[\bar{m}|B'] = 1$ into (9). Then,

$$\varpi(\bar{\theta}) = F - \bar{\theta}\eta(\bar{m}|\widehat{\theta}^P_{B'} = \bar{\theta})(P-c) = F - 0.5\bar{\theta}(P-c) \geq 0, \tag{A.37}$$

where the second equality follows from (A.36) and the inequality by the assumption of the case. Then, the sender has no incentives to deviate if $\widehat{\theta}^P_{B'} = \widehat{\theta}^P_{N'} = \bar{\theta}$ by Lemma A.6. At the same time, it follows from Lemma A.7 and (A.37) that

$$\forall \theta < \bar{\theta}, \varpi(\theta) > 0. \tag{A.38}$$

Consequently, by Lemma A.6 there are no possible cutoffs except for $\widehat{\theta}^P_{B'} = \widehat{\theta}^P_{N'} = \bar{\theta}$ such that the sender does not have an incentive to deviate.

Case 2: $F \in [\bar{\theta}\frac{P-c}{4-\kappa}, 0.5\bar{\theta}(P-c))$. We have

$$\begin{aligned} \varpi(0) &= F - 0 \cdot \eta(\bar{m}|\widehat{\theta}^P_{B'} = 0)(P-c) = F > 0, & \text{(A.39)} \\ \varpi(\bar{\theta}) &= F - \bar{\theta}\eta(\bar{m}|\widehat{\theta}^P_{B'} = \bar{\theta})(P-c) = F - 0.5\bar{\theta}(P-c) < 0, & \text{(A.40)} \end{aligned}$$

where the last inequality follows by the assumption of the case. Then, from Lemma A.7 and the intermediate value theorem it follows that there exists a unique cutoff value $0 < \widehat{\theta}^P_{B'} < \bar{\theta}$ such that $\varpi(\widehat{\theta}^P_{B'}) = 0$ (the necessary and sufficient condition for an interior cutoff by Lemma A.6). At the same time, the cutoff $\widehat{\theta}^P_{B'} = \bar{\theta}$ is impossible due to (A.40) and Lemma A.6, so that the existing interior cutoff is the only possible cutoff. Lemma A.6 also implies that the corresponding unique cutoff in state N' where the sender does not have incentives to deviate is then given by $\widehat{\theta}^P_{N'} = \min\left\{\bar{\theta}, 2\widehat{\theta}^P_{B'}\right\}$.

Let us show that in this case $\widehat{\theta}^P_{B'} \geq 0.5\bar{\theta}$. By Lemmas A.7 and A.9 it holds

$$\widehat{\theta}^P_{B'} \geq 0.5\bar{\theta} \Leftrightarrow \varpi(0.5\bar{\theta}) \geq 0. \tag{A.41}$$

From (9) we get

$$\eta(\bar{m}|\widehat{\theta}^P_{B'} = 0.5\bar{\theta}) = \frac{2}{4-\kappa} \tag{A.42}$$

so that

$$\varpi(0.5\bar{\theta}) = F - 0.5\bar{\theta}\frac{2}{4-\kappa}(P-c) \geq 0, \tag{A.43}$$

where the inequality follows from the assumption $F \in [\bar{\theta}\frac{P-c}{4-\kappa}, 0.5\bar{\theta}(P-c))$. By (A.41) and (A.43) it then follows that $\widehat{\theta}^{P}_{B'} \geq 0.5\bar{\theta}$.

Case 3: $F < \bar{\theta}\frac{P-c}{4-\kappa}$. From $\varpi(0) > 0$, $\varpi(\bar{\theta}) < 0$ and Lemmas A.7 and A.6 it follows that there are unique (interior) cutoffs $\widehat{\theta}^{P}_{B'}$ and $\widehat{\theta}^{P}_{N'}$ with $\widehat{\theta}^{P}_{N'} = \min\left\{\bar{\theta}, 2\widehat{\theta}^{P}_{B'}\right\}$. Finally, (A.41), the left equality in (A.43) and $F < \bar{\theta}\frac{P-c}{4-\kappa}$ result in $\widehat{\theta}^{P}_{B'} < 0.5\bar{\theta}$. ■

Proof of Proposition 3. Lemma A.8 shows that for any parameter values there exist unique cutoffs $\widehat{\theta}^{P}_{B'}$ and $\widehat{\theta}^{P}_{N'}$ such that the sender does not have an incentive to deviate once the receiver plays according to the prescribed equilibrium strategy. Besides, the receiver does not have an incentive to deviate from her prescribed strategy after $\varnothing$ by (2). Thus, to prove the claim of the proposition we need to show that the remaining receiver's incentive constraint

$$\eta(\bar{m}|\widehat{\theta}^{P}_{B'}) > \underline{\eta}$$

(ensuring investment after $\bar{m}$) is satisfied given the unique possible equilibrium cutoffs $\widehat{\theta}^{P}_{B'}$ and $\widehat{\theta}^{P}_{N'}$. By Lemma 5 we have

$$\begin{aligned}
\eta(\bar{m}) &= \frac{\Pr[\bar{m}|G']\kappa + \Pr[\bar{m}|N'](1-\kappa)}{(\Pr[\bar{m}|G'] + \Pr[\bar{m}|B'])\kappa + 2\Pr[\bar{m}|N'](1-\kappa)} \\
&\geq \frac{\Pr[\bar{m}|G']\kappa + \Pr[\bar{m}|N'](1-\kappa)}{2\Pr[\bar{m}|G']\kappa + 2\Pr[\bar{m}|N'](1-\kappa)} = 0.5 > \underline{\eta},
\end{aligned} \tag{A.44}$$

where the first inequality follows from the fact that $\Pr[\bar{m}|B'] \leq \Pr[\bar{m}|G']$ by Lemma 3, and the last inequality is by Assumption 1. Hence, the receiver always finds it optimal to invest after $\bar{m}$.

Thus, for any possible cutoffs $\widehat{\theta}^{P}_{B'}$ and $\widehat{\theta}^{P}_{N'}$ where the sender does not have incentives to deviate (which in turn exist for any parameter values by Lemma A.8) the receiver also does not have incentives to deviate. Moreover, the values of the cutoffs in any pooling equilibrium uniquely determine the sender's equilibrium strategy (by Definition 1), and the receiver's equilibrium beliefs and hence strategy. Since these cutoff values are unique by Lemma A.8, any given pooling equilibrium is essentially unique (i.e. all other pooling equilibria can be different only in terms of exogenous formulation of $\bar{m}$). This completes the proof. ■

A.5 Separating equilibrium

Lemma A.9 *If the strategies and beliefs are specified according to Definition 2, then the sender has no incentives to deviate if and only if the following two conditions hold:*

1) either $\widehat{\theta}^{S}_{B'} < \bar{\theta}$ *and* $U^{s}_{B'}(\widehat{\theta}^{S}_{B'}, \eta(\widetilde{m}), I) = 0$ *or* $\widehat{\theta}^{S}_{B'} = \bar{\theta}$ *and* $U^{s}_{B'}(\bar{\theta}, \eta(\widetilde{m}), I) \geq 0$.

2) $\widehat{\theta}^{S}_{N'} = \min\left\{\bar{\theta}, 2\widehat{\theta}^{S}_{B'}\right\}$.

Proof. Note that no type sending $\widetilde{m}$ in states B' or N' has an incentive to deviate to $\bar{m}$ (leading to $\eta(\bar{m}) = 1$) due to higher expected guilt. At the same time, all types in state

G' strictly prefer $\bar{m}$ over $\widetilde{m}$ by lexicographic preferences (Assumption 2). The subsequent proof is based on the same arguments as the proof of Lemma A.6 for the case of the pooling equilibrium. ■

Corollary A.2 *In any separating equilibrium,* $\widehat{\theta}^S_{B'} = \min\left\{\frac{F}{\eta(\widetilde{m})(P-c)}, \bar{\theta}\right\}$.

Proof. Note that $\min\left\{\frac{F}{\eta(\widetilde{m})(P-c)}, \bar{\theta}\right\} = \bar{\theta}$ if and only if $U^s_{B'}(\bar{\theta}, \eta(\widetilde{m}), I) \geq 0$. Then, the result follows from Lemma A.9. ■

Lemma A.10 *If the strategies and beliefs are specified as in Definition 2 and* $\widehat{\theta}^S_{N'} = \min\left\{\bar{\theta}, 2\widehat{\theta}^S_{B'}\right\}$*, then* $U^s_{B'}(\widehat{\theta}^S_{B'}, \eta(\widetilde{m}|\widehat{\theta}^S_{B'}), I)$ *is continuous and strictly decreasing in* $\widehat{\theta}^S_{B'}$ *on* $(0, \bar{\theta}]$.

Proof. As in the case of the pooling equilibrium (see (A.27)-(A.28)), $U^s_{B'}(\widehat{\theta}^S_{B'}, \eta(\widetilde{m}|\widehat{\theta}^S_{B'}), I)$ is continuous in $\widehat{\theta}^S_{B'}$, while

$$\frac{dU^s_{B'}(\widehat{\theta}^S_{B'}, \eta(\widetilde{m}|\widehat{\theta}^S_{B'}), I)}{d\widehat{\theta}^S_{B'}} = -(P-c)\widehat{\theta}^S_{B'}\left(\frac{\partial\eta(\widetilde{m}|\widehat{\theta}^S_{B'})}{\partial\widehat{\theta}^S_{B'}} + \frac{\eta(\widetilde{m}|\widehat{\theta}^S_{B'})}{\widehat{\theta}^S_{B'}}\right) \tag{A.45}$$

for any $\widehat{\theta}^S_{B'} \in (0, \bar{\theta}]$ except for $\widehat{\theta}^S_{B'} = 0.5\bar{\theta}$ and $\widehat{\theta}^S_{B'} = \bar{\theta}$. Consider the following possible cases given that $\widehat{\theta}^S_{N'} = \min\left\{\bar{\theta}, 2\widehat{\theta}^S_{B'}\right\}$.

Case 1: $\widehat{\theta}^S_{B'} \in (0, 0.5\bar{\theta})$, $\widehat{\theta}^S_{N'} = 2\widehat{\theta}^S_{B'}$.

Then, by (9) (substituting $\Pr[\widetilde{m}|G'] = 0$, $\Pr[\widetilde{m}|N'] = 2\widehat{\theta}^S_{B'}/\bar{\theta}$ and $\Pr[\widetilde{m}|B'] = \widehat{\theta}^S_{B'}/\bar{\theta}$)

$$\eta(\widetilde{m}|\widehat{\theta}^S_{B'}) = \frac{2(1-\kappa)}{4-3\kappa}, \tag{A.46}$$

i.e. is constant. Thus,

$$\frac{\partial\eta(\widetilde{m}|\widehat{\theta}^S_{B'})}{\partial\widehat{\theta}^S_{B'}} = 0 \tag{A.47}$$

that together with (A.45) leads to the claim for $\widehat{\theta}^S_{B'} \in (0, 0.5\bar{\theta})$.

Case 2: $\widehat{\theta}^S_{B'} \in (0.5\bar{\theta}, \bar{\theta})$, $\widehat{\theta}^S_{N'} = \bar{\theta}$.

In this case, $\Pr[\widetilde{m}|G'] = 0$, $\Pr[\widetilde{m}|N'] = 1$ and $\Pr[\widetilde{m}|B'] = \widehat{\theta}^S_{B'}/\bar{\theta}$ so that (9) implies

$$\eta(\widetilde{m}|\widehat{\theta}^S_{B'}) = \frac{(1-\kappa)\bar{\theta}}{2(1-\kappa)\bar{\theta} + \kappa\widehat{\theta}^S_{B'}}. \tag{A.48}$$

Then, it is possible to obtain a simple closed-form solution for the RHS of (A.45):

$$
\begin{aligned}
&\frac{\partial\eta(\widetilde{m}|\widehat{\theta}^S_{B'})}{\partial\widehat{\theta}^S_{B'}}+\frac{\eta(\widetilde{m}|\widehat{\theta}^S_{B'})}{\widehat{\theta}^S_{B'}} \\
=\;& -\frac{(1-\kappa)\kappa\bar{\theta}}{(\kappa\widehat{\theta}^S_{B'}+2(1-\kappa)\bar{\theta})^2}+\frac{(1-\kappa)\bar{\theta}}{\widehat{\theta}^S_{B'}(\kappa\widehat{\theta}^S_{B'}+2(1-\kappa)\bar{\theta})} \\
=\;& \frac{2(1-\kappa)^2\bar{\theta}^2}{\widehat{\theta}^S_{B'}(\kappa\widehat{\theta}^S_{B'}+2(1-\kappa)\bar{\theta})^2}>0. \qquad \text{(A.49)}
\end{aligned}
$$

By (A.49) and (A.45) the claim follows (for $\widehat{\theta}^S_{B'}\in(0.5\bar{\theta},\bar{\theta})$).

Finally, since $\eta(\widetilde{m}|\widehat{\theta}^S_{B'})$, and hence $U^s_{B'}(\widehat{\theta}^S_{B'},\eta(\widetilde{m}|\widehat{\theta}^S_{B'}),I)$, is continuous at any $\widehat{\theta}^S_{B'}\in(0,\bar{\theta}]$ the claim holds for the whole interval $(0,\bar{\theta}]$. ■

Lemma A.11 *If the strategies and beliefs are specified as in Definition 2, then for any given parameter values there always exist unique cutoffs $\widehat{\theta}^S_{B'}$ and $\widehat{\theta}^S_{N'}$ such that the sender does not have incentives to deviate. Moreover:*

1) If $F\geq\frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}$ then $\widehat{\theta}^S_{B'}=\widehat{\theta}^S_{N'}=\bar{\theta}$.

2) If $F\in\left[\frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa},\frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}\right)$ then $\widehat{\theta}^S_{B'}\in[0.5\bar{\theta},\bar{\theta})$ and $\widehat{\theta}^S_{N'}=\bar{\theta}$.

3) If $F<\frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}$ then $\widehat{\theta}^S_{B'}\in(0,0.5\bar{\theta})$ and $\widehat{\theta}^S_{N'}=2\widehat{\theta}^S_{B'}$.

Proof. The proof proceeds analogously to the case of the pooling equilibrium. For notational simplicity denote

$$\phi(\widehat{\theta}^S_{B'})=U^s_{B'}(\widehat{\theta}^S_{B'},\eta(\widetilde{m}|\widehat{\theta}^S_{B'}),I) \qquad \text{(A.50)}$$

with $\widehat{\theta}^S_{N'}=\min\left\{\bar{\theta},2\widehat{\theta}^P_{B'}\right\}$. Consider the cases listed in the lemma.

Case 1: $F\geq\frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}$. Consider the behavior of $\phi(\cdot)$ at the exterior point $\bar{\theta}$. We have

$$\eta(\widetilde{m}|\widehat{\theta}^S_{B'}=\bar{\theta})=\frac{1-\kappa}{2-\kappa}, \qquad \text{(A.51)}$$

which results from substituting $\Pr[\widetilde{m}|G']=0$ and $\Pr[\bar{m}|N']=\Pr[\bar{m}|B']=1$ into (9). Then,

$$\phi(\bar{\theta})=F-\bar{\theta}\eta(\widetilde{m}|\widehat{\theta}^S_{B'}=\bar{\theta})(P-c)=F-\bar{\theta}\frac{1-\kappa}{2-\kappa}(P-c)\geq 0, \qquad \text{(A.52)}$$

where the second equality follows from (A.51) and the inequality by the assumption of the case. Then, the sender has no incentives to deviate if $\widehat{\theta}^S_{B'}=\widehat{\theta}^S_{N'}=\bar{\theta}$ by Lemma A.9. At the same time, it follows from Lemma A.10 and (A.52) that

$$\forall\theta<\bar{\theta},\phi(\theta)>0. \qquad \text{(A.53)}$$

Consequently, by Lemma A.9 there are no possible cutoffs except for $\widehat{\theta}^S_{B'}=\widehat{\theta}^S_{N'}=\bar{\theta}$ such that the sender does not have incentives to deviate.

Case 2: $F \in \left[\frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}, \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}\right)$. We have

$$\begin{aligned}
\phi(0) &= F - 0 \cdot \eta(\widetilde{m}|\widehat{\theta}^S_{B'} = 0)(P-c) = F > 0, && \text{(A.54)}\\
\phi(\bar{\theta}) &= F - \bar{\theta}\eta(\widetilde{m}|\widehat{\theta}^S_{B'} = \bar{\theta})(P-c) = F - \bar{\theta}\frac{1-\kappa}{2-\kappa}(P-c) < 0, && \text{(A.55)}
\end{aligned}$$

where the second equality follows by (A.51), and the inequality follows by the assumption of the case. Then, from Lemma A.10 and the intermediate value theorem it follows that there exists a unique cutoff value $0 < \widehat{\theta}^S_{B'} < \bar{\theta}$ such that $\phi(\widehat{\theta}^S_{B'}) = 0$ (the necessary and sufficient condition for an interior cutoff by Lemma A.9). At the same time, the cutoff $\widehat{\theta}^S_{B'} = \bar{\theta}$ is impossible due to (A.55) and Lemma A.9, so that the existing interior cutoff is the only possible cutoff. Lemma A.9 also implies that the corresponding unique cutoff in state N' is then given by $\widehat{\theta}^S_{N'} = \min\left\{\bar{\theta}, 2\widehat{\theta}^P_{B'}\right\}$.

Let us show that in this case $\widehat{\theta}^S_{B'} \geq 0.5\bar{\theta}$. By Lemmas A.10 and A.9 it holds

$$\widehat{\theta}^S_{B'} \geq 0.5\bar{\theta} \Leftrightarrow \phi(0.5\bar{\theta}) \geq 0. \tag{A.56}$$

From (9) we get

$$\eta(\widetilde{m}|\widehat{\theta}^S_{B'} = 0.5\bar{\theta}) = \frac{2(1-\kappa)}{4-3\kappa} \tag{A.57}$$

so that

$$\phi(0.5\bar{\theta}) = F - 0.5\bar{\theta}\frac{2(1-\kappa)}{4-3\kappa}(P-c) \geq 0, \tag{A.58}$$

where the inequality follows from the assumption $F \in \left[\frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}, \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}\right)$. By (A.56) and (A.58) it then follows that $\widehat{\theta}^S_{B'} \geq 0.5\bar{\theta}$.

Case 3: $F < \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}$. From $\phi(0) > 0$, $\phi(\bar{\theta}) < 0$ and Lemmas A.10 and A.9 it follows that there are unique interior cutoffs $\widehat{\theta}^S_{B'}$ and $\widehat{\theta}^S_{N'}$ with $\widehat{\theta}^S_{N'} = \min\left\{\bar{\theta}, 2\widehat{\theta}^P_{B'}\right\}$. Finally, (A.56), the left equality in (A.58) and $F < \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}$ result in $\widehat{\theta}^S_{B'} < 0.5\bar{\theta}$. ■

Proof of Proposition 4. To show the claim of the proposition we need to find the range of parameters such that the receiver's incentive constraints are satisfied given the unique equilibrium cutoffs $\widehat{\theta}^S_{B'}$ and $\widehat{\theta}^S_{N'}$ where the sender does not have incentives to deviate, which always exist by Lemma A.11 (in which case the obtained separating equilibrium will be essentially unique given the unique values of the cutoffs). Clearly, since message $\bar{m}$ is sent by the sender only if he has indeed observed the good state of the world, it holds $\eta(\bar{m}) = 1 > \underline{\eta}$, so that the receiver always prefers to invest after $\bar{m}$. Let us consider the remaining incentive constraint which ensures investment after $\widetilde{m}$

$$\eta(\widetilde{m}|\widehat{\theta}^S_{B'}) > \underline{\eta} = -\frac{c}{P-c}. \tag{A.59}$$

We consider this constraint in three possible parameter cases according to Lemma A.11.

Case 1: $F \geq \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}$ and $\widehat{\theta}^S_{B'} = \widehat{\theta}^S_{N'} = \bar{\theta}$.

By (A.51) we have

$$\eta(\widetilde{m}|\widehat{\theta}^S_{B'} = \bar{\theta}) = \frac{1-\kappa}{2-\kappa} \tag{A.60}$$

so that

$$\eta(\widetilde{m}|\widehat{\theta}^S_{B'} = \bar{\theta}) > -\frac{c}{P-c} \Leftrightarrow \kappa < \frac{P+c}{P}. \tag{A.61}$$

Case 2: $F \in \left[\frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}, \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}\right)$, $\widehat{\theta}^S_{B'} \in [0.5\bar{\theta}, \bar{\theta})$ and $\widehat{\theta}^S_{N'} = \bar{\theta}$.

Substituting for $\eta(\widetilde{m}|\widehat{\theta}^S_{B'})$ given that $\widehat{\theta}^S_{N'} = \bar{\theta}$ we get

$$\eta(\widetilde{m}|\widehat{\theta}^S_{B'}) = \frac{\bar{\theta}(1-\kappa)}{(1-\kappa)2\bar{\theta} + \kappa\widehat{\theta}^S_{B'}}. \tag{A.62}$$

Since $\eta(\widetilde{m}|\widehat{\theta}^S_{B'})$ is differentiable on $(0.5\bar{\theta}, \bar{\theta})$, for this interval

$$\frac{\partial\eta(\widetilde{m}|\widehat{\theta}^S_{B'})}{\partial F} = \frac{\partial\eta(\widetilde{m}|\widehat{\theta}^S_{B'})}{\partial\widehat{\theta}^S_{B'}}\frac{\partial\widehat{\theta}^S_{B'}}{\partial F}. \tag{A.63}$$

The first term in the RHS is

$$\frac{\partial\eta(\widetilde{m}|\widehat{\theta}^S_{B'})}{\partial\widehat{\theta}^S_{B'}} = -\frac{\bar{\theta}\kappa(1-\kappa)}{\left(2\bar{\theta}(1-\kappa) + \kappa\widehat{\theta}^S_{B'}\right)^2} < 0. \tag{A.64}$$

Consider the second term. By the implicit function theorem and the fact that $\phi(\widehat{\theta}^S_{B'}) = 0$ (since $\widehat{\theta}^S_{B'}$ is interior by assumption) we have

$$\frac{\partial\widehat{\theta}^S_{B'}}{\partial F} = -\frac{\partial\phi/\partial F}{\partial\phi/\partial\widehat{\theta}^S_{B'}} = -\frac{1}{\partial\phi/\partial\widehat{\theta}^S_{B'}} > 0, \tag{A.65}$$

where the last inequality follows by Lemma A.10. Finally, (A.63)-(A.65) lead to

$$\frac{\partial\eta(\widetilde{m}|\widehat{\theta}^S_{B'})}{\partial F} < 0. \tag{A.66}$$

Given continuity of $\eta(\widetilde{m}|\widehat{\theta}^S_{B'})$, we obtain that it is strictly decreasing on $[0.5\bar{\theta}, \bar{\theta})$. Then, under considered range of parameters, $\eta(\widetilde{m}|\widehat{\theta}^S_{B'})$ obtains its highest value if $F = \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}$ so that, correspondingly, $\widehat{\theta}^S_{B'} = 0.5\bar{\theta}$ (see (A.58)). In this case

$$\eta(\widetilde{m}|\widehat{\theta}^S_{B'} = 0.5\bar{\theta}) = \frac{2(1-\kappa)}{4-3\kappa}. \tag{A.67}$$

Consequently, if

$$\begin{aligned} \frac{2(1-\kappa)}{4-3\kappa} &\leq \underline{\eta} \\ &\Leftrightarrow \kappa \geq \frac{2(P+c)}{2P+c}, \end{aligned} \tag{A.68}$$

then for any F in the considered parameter range the incentive constraint $\eta(\widetilde{m}|\widehat{\theta}^S_{B'}) > \underline{\eta}$ is violated.

At the same time, by (A.64) $\eta(\widetilde{m}|\widehat{\theta}^S_{B'})$ is bounded from below by its value at $\widehat{\theta}^S_{B'} = \bar{\theta}$. In this case,

$$\eta(\widetilde{m}|\widehat{\theta}^S_{B'} = \bar{\theta}) = \frac{1-\kappa}{2-\kappa}. \tag{A.69}$$

Consequently, if

$$\begin{aligned} \frac{1-\kappa}{2-\kappa} &> \underline{\eta} \\ &\Leftrightarrow \kappa < \frac{P+c}{P}, \end{aligned} \tag{A.70}$$

then for any F in the considered case the incentive constraint $\eta(\widetilde{m}|\widehat{\theta}^S_{B'}) > \underline{\eta}$ is satisfied.

Next, consider the case when $\kappa \in \left[\frac{P+c}{P}, \frac{2(P+c)}{2P+c}\right)$. Then,

$$\eta(\widetilde{m}|\widehat{\theta}^S_{B'} = 0.5\bar{\theta}) = \frac{2(1-\kappa)}{4-3\kappa} > \underline{\eta}, \tag{A.71}$$

$$\eta(\widetilde{m}|\widehat{\theta}^S_{B'} = \bar{\theta}) = \frac{1-\kappa}{2-\kappa} \leq \underline{\eta}. \tag{A.72}$$

This, together with (A.64) and the intermediate value theorem, implies that there exists a threshold value of $\widehat{\theta}^S_{B'} \in (0.5\bar{\theta}, \bar{\theta}]$ such that the incentive constraint binds, i.e. $\eta(\widetilde{m}|\widehat{\theta}^S_{B'}) = \underline{\eta}$. This equality yields

$$\begin{aligned} \eta(\widetilde{m}|\widehat{\theta}^S_{B'}) &= \frac{\bar{\theta}(1-\kappa)}{(1-\kappa)2\bar{\theta} + \kappa\widehat{\theta}^S_{B'}} = -\frac{c}{P-c}, \\ \widehat{\theta}^S_{B'} &= \frac{(P+c)(1-\kappa)\bar{\theta}}{-c\kappa}. \end{aligned} \tag{A.73}$$

Substituting this into $\phi(\widehat{\theta}^S_{B'})$ we get

$$\begin{aligned} \phi(\widehat{\theta}^S_{B'}) &= F - \frac{(P+c)(1-\kappa)\bar{\theta}}{-c\kappa}\left(-\frac{c}{P-c}\right)(P-c) \\ &= F - \frac{(P+c)(1-\kappa)\bar{\theta}}{\kappa}, \end{aligned} \tag{A.74}$$

which together with the cutoff condition $\phi(\widehat{\theta}^S_{B'}) = 0$ (see Lemma A.9) implies that the value of F which leads to $\eta(\widetilde{m}|\widehat{\theta}^S_{B'}) = \underline{\eta}$ is

$$F_{|\eta(\widetilde{m}|\widehat{\theta}^S_{B'})=\underline{\eta}} = \bar{\theta}\,(P+c)\,\frac{(1-\kappa)}{\kappa}. \tag{A.75}$$

Then, since $\eta(\widetilde{m}|\widehat{\theta}^S_{B'})$ decreases with F in the considered case by (A.66), the receiver's incentive constraint (A.59) is satisfied, given $\kappa \in \left[\frac{P+c}{P}, \frac{2(P+c)}{2P+c}\right)$, if and only if

$$F < \bar{\theta}\,(P+c)\,\frac{(1-\kappa)}{\kappa}. \tag{A.76}$$

Note also that for the considered range of κ it holds

$$\bar{\theta}\,(P+c)\,\frac{(1-\kappa)}{\kappa} \leq \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa},$$

i.e. (A.76) is a binding constraint in Case 2.

In sum, in Case 2 the receiver's incentive constraints are satisfied whenever $\kappa < \frac{P+c}{P}$ for any F or $\kappa \in \left[\frac{P+c}{P}, \frac{2(P+c)}{2P+c}\right)$ and $F < \bar{\theta}\,(P+c)\,\frac{(1-\kappa)}{\kappa}$.

Case 3: $F < \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}$, $\widehat{\theta}^S_{B'} \in (0, 0.5\bar{\theta})$ and $\widehat{\theta}^S_{N'} = 2\widehat{\theta}^S_{B'}$.

Substituting $\Pr[\widetilde{m}|G'] = 0$, $\Pr[\widetilde{m}|N'] = 2\widehat{\theta}^S_{B'}/\bar{\theta}$ and $\Pr[\widetilde{m}|B'] = \widehat{\theta}^S_{B'}/\bar{\theta}$ into (9) we get

$$\eta(\widetilde{m}|\widehat{\theta}^S_{B'}) = \frac{2(1-\kappa)}{4-3\kappa}. \tag{A.77}$$

Then,

$$\begin{aligned} \frac{2(1-\kappa)}{4-3\kappa} \quad &> \quad \underline{\eta} = -\frac{c}{P-c} \\ &\Leftrightarrow \quad 0 < \kappa < \frac{2(P+c)}{2P+c}, \end{aligned} \tag{A.78}$$

determining the parameter range where the incentive constraint (A.59) holds.

Merging Case 1, Case 2 and Case 3 together, we obtain that an (essentially unique) separating equilibrium exists if and only if either of the following holds:

Case 1:

$$F \geq \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa} \wedge \kappa < \frac{P+c}{P};$$

Case 2:

$$\left(F \in \left[\frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}, \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}\right) \wedge \kappa < \frac{P+c}{P}\right) \text{ or}$$
$$\left(F \in \left[\frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}, \bar{\theta}\,(P+c)\,\frac{(1-\kappa)}{\kappa}\right) \wedge \kappa \in \left[\frac{P+c}{P}, \frac{2(P+c)}{2P+c}\right)\right);$$

Case 3:

$$F \in \left(0, \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}\right) \wedge \kappa < \frac{2(P+c)}{2P+c}.$$

This is equivalent to the statement of the proposition. ■

Proof of Lemma 6. Given that $\widehat{\theta}^P_{N'} = \min\left\{\bar{\theta}, 2\widehat{\theta}^P_{B'}\right\}$ and $\widehat{\theta}^S_{N'} = \min\left\{\bar{\theta}, 2\widehat{\theta}^S_{B'}\right\}$ (by Lemmas A.6 and A.9), it is sufficient to show the claim for $\widehat{\theta}^S_{B'}$ and $\widehat{\theta}^P_{B'}$. We have that in a pooling equilibrium for any $\widehat{\theta}^P_{B'}$

$$\eta(\bar{m}|\widehat{\theta}^P_{B'}) \geq 0.5 \tag{A.79}$$

(see (A.44)). At the same time, in a separating equilibrium

$$\begin{aligned}\eta(\widetilde{m}|\widehat{\theta}^S_{B'}) &= \Pr[G|\widetilde{m}] = \Pr[G|N' \cap \widetilde{m}] \Pr[N'|\widetilde{m}] + \Pr[G|B' \cap \widetilde{m}] \Pr[B'|\widetilde{m}] \\ &= 0.5 \Pr[N'|\widetilde{m}] < 0.5, \end{aligned} \tag{A.80}$$

because at least some types in state B' send $\widetilde{m}$ so that $\Pr[N'|\widetilde{m}] < 1$. (A.79) and (A.80) yield

$$\eta(\bar{m}|\widehat{\theta}^P_{B'}) > \eta(\widetilde{m}|\widehat{\theta}^S_{B'}) \tag{A.81}$$

so that

$$U^s_{B'}(\widehat{\theta}^P_{B'}, \eta(\widetilde{m}|\widehat{\theta}^S_{B'}), I) > U^s_{B'}(\widehat{\theta}^P_{B'}, \eta(\bar{m}|\widehat{\theta}^P_{B'}), I). \tag{A.82}$$

If $\widehat{\theta}^P_{B'} < \bar{\theta}$ so that $U^s_{B'}(\widehat{\theta}^P_{B'}, \eta(\bar{m}|\widehat{\theta}^P_{B'}), I) = 0$ by Lemma A.6, then (A.82) implies

$$U^s_{B'}(\widehat{\theta}^P_{B'}, \eta(\widetilde{m}|\widehat{\theta}^S_{B'}), I) > 0. \tag{A.83}$$

Then, $\widehat{\theta}^S_{B'} > \widehat{\theta}^P_{B'}$ by Lemma A.9 and the fact that $U^s_{B'}(\theta, \eta(\widetilde{m}|\widehat{\theta}^S_{B'}), I)$ is decreasing in the first argument.

In the other case, if $\widehat{\theta}^P_{B'} = \bar{\theta}$, by (A.82) and Lemma A.6

$$U^s_{B'}(\bar{\theta}, \eta(\widetilde{m}|\widehat{\theta}^S_{B'}), I) > 0. \tag{A.84}$$

Consequently, all types in state B' prefer to send $\widetilde{m}$ in a separating equilibrium and the only possible equilibrium cutoff is $\widehat{\theta}^S_{B'} = \bar{\theta}$. ■

A.6 Other equilibria under guilt aversion

Lemma A.12 *If any two equilibria are characterized by the same persuasiveness of investment-inducing messages in states $i^s \in \{B', N'\}$ (in the sense of Lemma A.1), then these equilibria are payoff-equivalent.*

Proof. Denote by $\widehat{\eta}$ the persuasiveness of messages used in states $i^s \in \{B', N'\}$ (unique by Lemma A.1). We need to show that, if in any two equilibria $\widehat{\eta}$ is the same, then all sender types in all states and the receiver have the same distribution of payoffs in both equilibria.

Claim 1. *The cutoffs* $\widehat{\theta}_{B'}$ *and* $\widehat{\theta}_{N'}$ *do not differ between the equilibria.*

Proof. Assume the opposite by contradiction. Then, given that $\widehat{\theta}_{N'}$ is uniquely determined by $\widehat{\theta}_{B'}$ by Lemma A.5, at least in one equilibrium it must hold $\widehat{\theta}_{B'} < \bar{\theta}$ (denote this cutoff as $\widehat{\theta}_{B'1}$). Then, by Lemma A.4 it should hold

$$U^s_{B'}(\widehat{\theta}_{B'1}, \widehat{\eta}, I) = 0,$$

which gives

$$\begin{aligned} \forall\theta &< \widehat{\theta}_{B'1}, U^s_{B'}(\theta, \widehat{\eta}, I) > 0, \\ \forall\theta &> \widehat{\theta}_{B'1}, U^s_{B'}(\theta, \widehat{\eta}, I) < 0. \end{aligned}$$

Then, by Lemma A.4 no equilibrium can have a cutoff in state B' different from $\widehat{\theta}_{B'1}$ given $\widehat{\eta}$, which yields a contradiction.

Claim 2. *The equilibria are payoff-equivalent for the receiver.*

Proof. The receiver's distribution of payoffs for given sender type θ and information state i^s depends only on the receiver's action and the state i^s (since the distribution of the state of the world is uniquely determined by i^s). At the same time, the correspondence of receiver's actions to sender types (except for the zero measured cutoff types) in each information state i^s is uniquely determined by the cutoff in this state (by Lemmas 4 and A.2). Since the cutoffs do not differ between the equilibria, the claim holds.

Claim 3. *The equilibria are payoff-equivalent for the sender.*

Consider any given sender's type θ in an information state i^s. Note first that the cutoffs $\widehat{\theta}^P_{i^s}$ for $i^s \in \{B', N'\}$ are the same between the equilibria by Claim 1. Then, we can have the following cases:

- If $i^s \in \{B', N'\}$ while $\theta > \widehat{\theta}^P_{i^s}$, then by Lemma 4 the sender obtains 0 in both equilibria.

- If $i^s \in \{B', N'\}$ while $\theta < \widehat{\theta}^P_{i^s}$, then by Lemma 4 and (10) the sender obtains the same expected payoff of $F - \theta\lambda_{i^s}\widehat{\eta}(P - c)$ in both equilibria.

- If $i^s \in \{B', N'\}$ while $\theta = \widehat{\theta}^P_{i^s} < \bar{\theta}$, then the sender must be indifferent between 0 and sending an investment-inducing message, and thus expects 0 in either equilibrium.

- If $i^s \in \{B', N'\}$ while $\theta = \widehat{\theta}^P_{i^s} = \bar{\theta}$, then the sender is either again indifferent between 0 and sending an investment-inducing message, or has a strictly positive utility from inducing investment in both equilibria obtaining the same expected payoff of $F - \theta\lambda_{i^s}\widehat{\eta}(P - c)$.

- If $i^s = G'$, then by Lemma 3 the sender always obtains F in both equilibria.

Hence, any given sender's type in any information state has the same expected utility in both equilibria. ■

Proof of Proposition 5.

Claim 1. *There can exist only two types of equilibria:*

- *where all types in state G' pool with types in the other states (Type 1)*;
- *where no type in state G' pools with types in the other states (Type 2).*

Proof. The claim is equivalent to showing that there exists no equilibrium where some types in state G' pool with types in the other states (i.e. send a message which is also chosen by some types in states B' or N'), and some types in state G' separate from types in the other states (i.e. do not send such a message). Assume by contradiction that this is the case. By Lemma 3, any separating type in state G' must also send a message. Hence, there exists at least one message m' sent by such types such that $\eta(m') = 1$. Then, this must hold for any m sent in state G' since otherwise the types sending messages with a lower persuasiveness in state G' would like to deviate to m' by lexicographic preferences (Assumption 2).

Next, note that by Lemma A.1 all investment-inducing messages sent in states B' and N' should have the same persuasiveness η'. Since at least some types in states B' and N' choose messages sent in state G' by assumption (and hence induce $\eta = 1$ by the previous argument), we must have $\eta' = 1$. This implies that all types in states B' and N' who send an investment-inducing message must pool at message(s) sent by types in state G', while any such message must be sent by a zero measure of types in states B' and N' (to ensure $\eta' = 1$). Thus, denoting the set of all messages inducing investment and sent by types in states B' and N' by $\Upsilon_{I,\{B',N'\}}$, we have

$$\forall m \in \Upsilon_{I,\{B',N'\}}\colon \ \Pr[m|\neg G'] = 0. \tag{A.85}$$

Denote the event that $m \in \Upsilon_{I,\{B',N'\}}$ as $\widetilde{\Upsilon}_{I,\{B',N'\}}$. Denote the set of types in state $i^s \in \{B', N'\}$ who send an investment-inducing message by Θ_{I,i^s}. Denote the event that $\theta \in \Theta_{>,i^s}$ in the sender's information state $i^s \in \{B', N'\}$ by $\widetilde{\Theta}_{I,\{B',N'\}}$. Then,

$$\begin{aligned}
\Pr[G|\widetilde{\Theta}_{I,\{B',N'\}}] &= \Pr[G|\widetilde{\Upsilon}_{I,\{B',N'\}} \cap \neg G'] && \text{(A.86)}\\
&= \sum_{m \in \Upsilon_{I,\{B',N'\}}} \Pr[G|m \cap \neg G'] \Pr[m|\neg G'] && \text{(A.87)}\\
&= 0, && \text{(A.88)}
\end{aligned}$$

where the first equality is by construction of the events, the second equality follows by the law of total probability and the countability of the message space, and the last equality follows from (A.85). Yet, $\Pr[G|\widetilde{\Theta}_{I,\{B',N'\}}] = 0$ contradicts to the fact that the good state has a positive probability in case of $i^s = N'$ while a positive measure of types in this informational state induce investment by Lemma 4. Hence, the constructed equilibrium does not exist.

Claim 2. *If an equilibrium is of Type 1 in the sense of Claim 1, then it is payoff-equivalent to a pooling equilibrium.*

Proof. Consider an equilibrium ζ of Type 1, where all types in state G' pool with types in the other states. By Lemma A.1 all investment-inducing messages in states B' and N' have the same persuasiveness η'. Then, all types in state G' must induce the same persuasiveness as well. Indeed, assume by contradiction that this is not the case, i.e. there exists at least one type θ' in state G' who sends a message m' such that $\eta(m') \neq \eta'$. Then, at least some types in states B' and N' must send m' as well since the equilibrium is of Type 1 by assumption. Yet, this contradicts Lemma A.1 due to $\eta(m') \neq \eta'$.

Next, denote the equilibrium cutoffs by $\widehat{\theta}^x_{B'}$ and $\widehat{\theta}^x_{N'}$ (in the sense of Lemma 4). Consider a messaging strategy of the sender denoted by σ such that all types in state $i^s \in \{B', N'\}$ with $\theta \leq \widehat{\theta}^x_{i^s}$ and all types in state G' pool on the same message $\bar{m}^x$, and the remaining types (if any) refrain from advice. Let us show that if such strategy is played in equilibrium, then $\eta(\bar{m}^x) = \eta'$.

Indeed, denote by Θ_{i^s} the set of all types in state $i^s \in \{B', N', G'\}$ inducing investment by sending a message in equilibrium ζ. Denote the event that $\theta \in \Theta_{i^s}$ in the sender's information state i^s by $\widetilde{\Theta}$. Denote the set of all investment-inducing messages in equilibrium ζ by Υ_I and the event that $m \in \Upsilon_I$ in this equilibrium as $\widetilde{\Upsilon_I}$. Note that in equilibrium ζ the event $\widetilde{\Theta}$ occurs if and only if the event $\widetilde{\Upsilon_I}$ occurs so that

$$\Pr[G|\widetilde{\Theta}] = \Pr[G|\widetilde{\Upsilon_I}]. \tag{A.89}$$

At the same time, we have

$$\begin{aligned} \Pr[G|\overline{m}^x] &= \Pr[G|\widetilde{\Theta}] = \Pr[G|\widetilde{\Upsilon_I}] = \sum_{m \in \Upsilon_I} \Pr[G|m \cap \widetilde{\Upsilon_I}] \Pr[m|\widetilde{\Upsilon_I}] \\ &= \sum_{m \in \Upsilon_I} \Pr[G|m] \Pr[m|\widetilde{\Upsilon_I}] = \sum_{m \in \Upsilon_I} \eta' \Pr[m|\widetilde{\Upsilon_I}] = \eta', \end{aligned} \tag{A.90}$$

where the first equality is by construction of $\bar{m}^x$, the second equality is by (A.89), the third equality is by the law of total probability and the fact that the message set is countable, the fourth equality is by the fact that any investment-inducing message m' is sent if and only if $m' \in \Upsilon_I$ by construction, and the fifth equality is by the fact that all investment-inducing messages have the same persuasiveness η' as shown above.

Next, consider a putative equilibrium ζ' where the sender plays σ, the receiver responds to $\bar{m}^x$ with investment, and to $\varnothing$ with investment if and only if $\eta(\varnothing) > \underline{\eta}$, while all out-of-equilibrium messages induce receiver's beliefs equal to $\eta(\bar{m}^x)$. Let us show that this is indeed an equilibrium. First, note that all types in state G' clearly find it optimal to send $\bar{m}^x$. Consider types in states B' and N'. For any $i^s \in \{B', N'\}$ it must hold

$$\begin{aligned} \forall\theta &\leq \widehat{\theta}^x_{i^s} : U^s_{i^s}(\theta, \eta\prime, I) \geq 0, \\ \forall\theta &> \widehat{\theta}^x_{i^s} : U^s_{i^s}(\theta, \eta\prime, I) < 0 \end{aligned}$$

since otherwise the sender's incentive constraints in equilibrium ζ would be violated (by Lemma 4). Consequently, given (A.90), the incentive constraints of the sender in states B' and N' in equilibrium ζ' are also satisfied. Besides, the receiver would prefer to invest after $\overline{m}^x$ (since the level of persuasiveness $\Pr[G|\overline{m}^x] = \eta'$ must be above $\underline{\eta}$ as otherwise η' would not induce investment in equilibrium ζ). Consequently, ζ' constitutes an existing pooling equilibrium. Moreover, it is payoff-equivalent to equilibrium ζ by Lemma A.12 since both ζ and ζ' are characterized by the same persuasiveness of investment-inducing messages in states $i^s \in \{B', N'\}$.

Claim 3. *If an equilibrium is of Type 2 in the sense of Claim 1, then it is payoff-equivalent to a separating equilibrium.*

Proof. Consider an equilibrium ζ of Type 2, where all types in state G' separate from the types in the other states. Again, by Lemma A.1 all investment-inducing messages in states B' and N' should have the same persuasiveness η', while all types in state G' must induce $\eta = 1$ since the equilibrium is assumed to be of Type 2.

Consider a messaging strategy of the sender denoted by σ' such that all types in state $i^s \in \{B', N'\}$ with $\theta \leq \widehat{\theta}^x_{i^s}$ pool on the same message $\widetilde{m}^x$, the remaining types in these states (if any) refrain from advice, and types in state G' separate with another message $\bar{m}^x$. By the same argument as in Case 1 (redefining Θ_{i^s} as the set of types sending an investment-inducing message in states B' and N', and Υ_I as the set of investment-inducing messages in states B' and N'), we obtain $\eta(\widetilde{m}^x) = \eta'$. Thus, strategy σ' leads to the same persuasiveness of investment-inducing messages as in equilibrium ζ. The remaining argument is analogous to Case 1 and is omitted.

The statement of the proposition follows jointly from Claims 1-3. ■

A.7 Welfare comparison

Lemma A.13 *Both $\widehat{\theta}^P_{B'}$ and $\widehat{\theta}^S_{B'}$ are continuous in $F \in (0, \infty)$ and $\kappa \in (0, 1)$.*

Proof.

Claim 1. *$\widehat{\theta}^P_{B'}$ is continuous in $F \in (0, \infty)$.*

Proof. Step 1. Consider $\widehat{\theta}^P_{B'}$ and its continuity in F. Note that $\eta(\bar{m}|\widehat{\theta}^P_{B'})$ is continuously differentiable in $\widehat{\theta}^P_{B'}$ on both intervals $(0, 0.5\theta)$ and $(0.5\bar{\theta}, \bar{\theta})$ by (A.29) and (A.33), respectively. Hence, $\varpi(\widehat{\theta}^P_{B'})$ is continuously differentiable in both F and $\widehat{\theta}^P_{B'}$ once $\widehat{\theta}^P_{B'}$ belongs to these intervals. Consequently, by the implicit function theorem (given the equilibrium condition $\varpi(\widehat{\theta}^P_{B'}) = 0$ for $\widehat{\theta}^P_{B'} < \bar{\theta}$ by Lemma A.6), $\widehat{\theta}^P_{B'}$ is continuous in F if $\widehat{\theta}^P_{B'} \in (0, 0.5\theta) \cup (0.5\bar{\theta}, \bar{\theta})$, which by Lemma A.8 holds if and only if

$$F \in \left(0, \bar{\theta}\frac{P-c}{4-\kappa}\right) \cup \left(\bar{\theta}\frac{P-c}{4-\kappa}, 0.5\bar{\theta}(P-c)\right).$$

Hence, we are left to show that $\widehat{\theta}^P_{B'}$ is continuous in F at $\bar{\theta}\frac{P-c}{4-\kappa}$ and $0.5\bar{\theta}(P-c)$, which is done in Steps 2 and 3, respectively.

Step 2. Let us show that $\widehat{\theta}^P_{B'}(F)$ is continuous at $F = \bar{\theta}\frac{P-c}{4-\kappa}$. By Lemma A.8 together with (A.29) and (A.33) we obtain:

$$\eta(\bar{m}|\widehat{\theta}^P_{B'}) = \begin{cases} \eta'(\widehat{\theta}^P_{B'}) \equiv \frac{2\widehat{\theta}^P_{B'}(1-\kappa)+\kappa\bar{\theta}}{\widehat{\theta}^P_{B'}(4-3\kappa)+\kappa\bar{\theta}} \text{ if } F < \bar{\theta}\frac{P-c}{4-\kappa}, \\ \eta''(\widehat{\theta}^P_{B'}) \equiv \frac{\bar{\theta}}{\kappa\widehat{\theta}^P_{B'}+(2-\kappa)\bar{\theta}} \text{ if } F \geq \bar{\theta}\frac{P-c}{4-\kappa}. \end{cases} \tag{A.91}$$

Furthermore, by Corollary A.1 and the fact that $\widehat{\theta}^P_{B'} < \bar{\theta}$ at $F = \bar{\theta}\frac{P-c}{4-\kappa}$ by Lemma A.8, we have

$$\widehat{\theta}^P_{B'} = \frac{F}{\eta(\bar{m}|\widehat{\theta}^P_{B'})(P-c)}. \tag{A.92}$$

Given (A.91), this can be rewritten as

$$\widehat{\theta}^P_{B'} = \begin{cases} \frac{F}{\eta'(\widehat{\theta}^P_{B'})(P-c)} \text{ if } F < \bar{\theta}\frac{P-c}{4-\kappa}, \\ \frac{F}{\eta''(\widehat{\theta}^P_{B'})(P-c)} \text{ if } F \geq \bar{\theta}\frac{P-c}{4-\kappa}. \end{cases} \tag{A.93}$$

Then, denoting $\overrightarrow{\theta}^- \equiv \lim_{F\to\bar{\theta}\frac{P-c}{4-\kappa}^-} \widehat{\theta}^P_{B'}$ we obtain

$$\begin{aligned} \overrightarrow{\theta}^- &= \frac{\bar{\theta}\frac{P-c}{4-\kappa}}{\lim_{F\to\bar{\theta}\frac{P-c}{4-\kappa}^-} \eta'(\widehat{\theta}^P_{B'})(P-c)} \\ &= \frac{\bar{\theta}\frac{P-c}{4-\kappa}}{\frac{2\overrightarrow{\theta}^-(1-\kappa)+\kappa\bar{\theta}}{\overrightarrow{\theta}^-(4-3\kappa)+\kappa\bar{\theta}}(P-c)}. \end{aligned} \tag{A.94}$$

Solving for $\overrightarrow{\theta}^-$ yields the only positive solution $\overrightarrow{\theta}^- = 0.5\bar{\theta}$.

Similarly, denoting $\overrightarrow{\theta}^+ \equiv \lim_{F\to\bar{\theta}\frac{P-c}{4-\kappa}^+} \widehat{\theta}^P_{B'}$ we obtain by (A.93)

$$\begin{aligned} \overrightarrow{\theta}^+ &= \frac{\bar{\theta}\frac{P-c}{4-\kappa}}{\lim_{F\to\bar{\theta}\frac{P-c}{4-\kappa}^+} \eta''(\widehat{\theta}^P_{B'})(P-c)} \\ &= \frac{\bar{\theta}\frac{P-c}{4-\kappa}}{\frac{\bar{\theta}}{\kappa\overrightarrow{\theta}^++(2-\kappa)\bar{\theta}}(P-c)}, \end{aligned} \tag{A.95}$$

which again yields $\overrightarrow{\theta}^+ = 0.5\bar{\theta}$. (A.94) and (A.95) together imply

$$\lim_{F\to\bar{\theta}\frac{P-c}{4-\kappa}^-} \widehat{\theta}^P_{B'} = \lim_{F\to\bar{\theta}\frac{P-c}{4-\kappa}^+} \widehat{\theta}^P_{B'}.$$

This proves that $\widehat{\theta}^P_{B'}(F)$ is continuous at $F = \bar{\theta}\frac{P-c}{4-\kappa}$.

Step 3. Let us show that $\widehat{\theta}^P_{B'}(F)$ is continuous at $F = 0.5\bar{\theta}(P-c)$.

By Corollary A.1 and A.8

$$\widehat{\theta}^P_{B'} = \begin{cases} \frac{F}{\eta''(\widehat{\theta}^P_{B'})(P-c)} \text{if } F \in [\bar{\theta}\frac{P-c}{4-\kappa}, 0.5\bar{\theta}(P-c)) \\ \bar{\theta} \text{ if } F \geq 0.5\bar{\theta}(P-c) \end{cases} . \tag{A.96}$$

Note that by Lemma A.8, the condition $\widehat{\theta}^P_{B'} < \bar{\theta}$ and hence (A.92) still hold for $F < 0.5\bar{\theta}(P-c)$. Then, denoting $\overrightarrow{\theta}^- \equiv \lim_{0.5\bar{\theta}(P-c)^-} \widehat{\theta}^P_{B'}$ we obtain

$$\begin{aligned} \overrightarrow{\theta}^- &= \frac{0.5\bar{\theta}(P-c)}{\lim_{F\to 0.5\bar{\theta}(P-c)^-} \eta''(\widehat{\theta}^P_{B'})(P-c)} \\ &= \frac{0.5\bar{\theta}(P-c)}{\frac{\bar{\theta}}{\kappa\overrightarrow{\theta}^- + (2-\kappa)\bar{\theta}}(P-c)}. \end{aligned} \tag{A.97}$$

Solving for $\overrightarrow{\theta}^-$ yields the only positive solution $\overrightarrow{\theta}^- = \bar{\theta}$. Hence,

$$\lim_{0.5\bar{\theta}(P-c)^-} \widehat{\theta}^P_{B'} = \bar{\theta} = \lim_{0.5\bar{\theta}(P-c)^+} \widehat{\theta}^P_{B'},$$

where the last equality is by (A.96). Consequently, $\widehat{\theta}^P_{B'}(F)$ is continuous at $F = 0.5\bar{\theta}(P-c)$.

Claim 2. $\widehat{\theta}^P_{B'}$ *is continuous in* $\kappa \in (0,1)$.

Note that the claim follows trivially if $F \geq 0.5\bar{\theta}(P-c)$, since then $\widehat{\theta}^P_{B'} = \bar{\theta}$ for any κ by Lemma A.8.

Consider $F < 0.5\bar{\theta}(P-c)$ so that $\widehat{\theta}^P_{B'} < \bar{\theta}$. Note that $\eta(\bar{m}|\widehat{\theta}^P_{B'})$ is continuously differentiable in $\widehat{\theta}^P_{B'}$ and κ for $\widehat{\theta}^P_{B'} \in (0, 0.5\theta) \cup (0.5\bar{\theta}, \bar{\theta})$ and $\kappa \in (0,1)$ by (A.29) and (A.33), respectively. The same holds for $\varpi(\widehat{\theta}^P_{B'})$. Consequently, by the implicit function theorem (given the equilibrium condition $\varpi(\widehat{\theta}^P_{B'}) = 0$ for $\widehat{\theta}^P_{B'} < \bar{\theta}$ by Lemma A.6), $\widehat{\theta}^P_{B'}$ is continuous in κ if $\widehat{\theta}^P_{B'} \in (0, 0.5\theta) \cup (0.5\bar{\theta}, \bar{\theta})$, which by Lemma A.8 holds if and only if

$$F \in \left(0, \bar{\theta}\frac{P-c}{4-\kappa}\right) \cup \left(\bar{\theta}\frac{P-c}{4-\kappa}, 0.5\bar{\theta}(P-c)\right).$$

Hence, we are left to show that $\widehat{\theta}^P_{B'}$ is continuous in κ if $F = \bar{\theta}\frac{P-c}{4-\kappa} \iff \kappa = (4F - \bar{\theta}(P-c))/F$. This follows by the same arguments as in Step 2 in the proof of Claim 1.

Claim 3. $\widehat{\theta}^S_{B'}$ *is continuous in* $F \in (0,\infty)$ *and* $\kappa \in (0,1)$.

Proof. The proof proceeds by the same arguments as the proofs of Claims 1 and 2, and hence is omitted. ■

Corollary A.3 *Both* $\widehat{\theta}^P_{N'}$ *and* $\widehat{\theta}^S_{N'}$ *are continuous in* $F \in (0,\infty)$ *and* $\kappa \in (0,1)$.

Proof. The claim immediately follows from Proposition A.13 and the fact that $\widehat{\theta}_{N'} = \min\left\{\bar{\theta}, 2\widehat{\theta}_{B'}\right\}$ by Lemma A.5. ■

Lemma A.14 *If at least some sender types refrain from advice in equilibrium, then:*

(i) the receiver invests conditional on $\varnothing$ *in the pooling equilibrium if and only if* $\kappa < \frac{P+c}{P}$ *and* $F < \widetilde{F}^P$ *where* $\widetilde{F}^P \in \left(0, \bar{\theta}\frac{P-c}{4-\kappa}\right)$ *is some threshold value.*

(ii) the receiver invests conditional on $\varnothing$ *in the separating equilibrium if and only if* $\kappa < \frac{P+c}{P}$ *and* $F < \widetilde{F}^S$ *where* $\widetilde{F}^S \in \left(0, \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}\right)$ *is some threshold value.*

Proof. (i) First, note that the receiver never invests after $\varnothing$ (if on the equilibrium path) if $\widehat{\theta}^P_{N'} = \bar{\theta}$, since then only types in state B' refrain in equilibrium so that $\eta(\varnothing) = 0$. Consider the remaining case $\widehat{\theta}^P_{N'} < \bar{\theta}$. Note that the receiver invests after $\varnothing$ if and only if

$$\eta(\varnothing|\widehat{\theta}^P_{B'}) > \underline{\eta}.$$

By Lemma 5

$$\eta(\varnothing) = \frac{\Pr[\varnothing|G']\kappa + \Pr[\varnothing|N'](1-\kappa)}{(\Pr[\varnothing|G'] + \Pr[\varnothing|B'])\kappa + 2\Pr[\varnothing|N'](1-\kappa)}. \tag{A.98}$$

Substituting $\Pr[\varnothing|G'] = 0$, $\Pr[\varnothing|N'] = (\bar{\theta} - \widehat{\theta}^P_{N'})/\bar{\theta}$, $\Pr[\varnothing|B'] = (\bar{\theta} - \widehat{\theta}^P_{B'})/\bar{\theta}$, and $\widehat{\theta}^P_{N'} = 2\widehat{\theta}^P_{B'}$ (due to Lemma A.5 and $\widehat{\theta}^P_{N'} < \bar{\theta}$) into (A.98) yields

$$\eta(\varnothing|\widehat{\theta}^P_{B'}) = \frac{(1-\kappa)(\bar{\theta} - 2\widehat{\theta}^P_{B'})}{(2-\kappa)\bar{\theta} - (4-3\kappa)\widehat{\theta}^P_{B'}}. \tag{A.99}$$

Next, let us show that then

$$\frac{\partial \eta(\varnothing|\widehat{\theta}^P_{B'})}{\partial F} < 0. \tag{A.100}$$

Since $\eta(\varnothing|\widehat{\theta}^P_{B'})$ depends on F only through $\widehat{\theta}^P_{B'}$, we have

$$\frac{\partial \eta(\varnothing|\widehat{\theta}^P_{B'})}{\partial F} = \frac{\partial \eta(\varnothing|\widehat{\theta}^P_{B'})}{\partial \widehat{\theta}^P_{B'}} \frac{\partial \widehat{\theta}^P_{B'}}{\partial F}. \tag{A.101}$$

The first term in the RHS is (differentiating (A.99))

$$\frac{\partial \eta(\varnothing|\widehat{\theta}^P_{B'})}{\partial \widehat{\theta}^P_{B'}} = -\frac{\bar{\theta}\kappa(1-\kappa)}{\left(\bar{\theta}(2-\kappa) - \widehat{\theta}^P_{B'}(4-3\kappa)\right)^2} < 0. \tag{A.102}$$

Consider the second term. By the implicit function theorem and the fact that $\varpi(\widehat{\theta}^P_{B'}) = 0$ (since $\widehat{\theta}^P_{B'}$ is interior in the considered case) we have

$$\frac{\partial \widehat{\theta}^P_{B'}}{\partial F} = -\frac{\partial \varpi / \partial F}{\partial \varpi / \partial \widehat{\theta}^P_{B'}} = -\frac{1}{\partial \varpi / \partial \widehat{\theta}^P_{B'}} > 0, \tag{A.103}$$

where the last inequality follows by Lemma A.7 and the fact that ϖ is differentiable in $\widehat{\theta}^P_{B'}$ on $(0, 0.5\bar{\theta})$. Finally, (A.101)-(A.103) lead to (A.100).

Let us consider the limit of $\eta(\varnothing|\widehat{\theta}^P_{B'})$ as F goes to 0 (note that the case restriction $\widehat{\theta}^P_{N'} < \bar{\theta}$ would then still hold by Lemma A.8). First, we have

$$\lim_{F\to 0} \widehat{\theta}^P_{B'} = \lim_{F\to 0} \frac{F}{\eta(m_{G'}|\widehat{\theta}^P_{B'})(P-c)} = 0, \tag{A.104}$$

where the first equality follows from Corollary A.1 and the last equality follows from the fact that $\eta(m_{G'}|\widehat{\theta}^P_{B'}) \geq 0.5$ by (A.44) and hence bounded from 0. Consequently,

$$\lim_{F\to 0} \eta(\varnothing|\widehat{\theta}^P_{B'}) = \lim_{\widehat{\theta}^P_{B'}\to 0} \frac{(1-\kappa)(\bar{\theta} - 2\widehat{\theta}^P_{B'})}{(2-\kappa)\bar{\theta} - (4-3\kappa)\widehat{\theta}^P_{B'}} = \frac{1-\kappa}{2-\kappa}. \tag{A.105}$$

Expressions (A.100) and (A.105) yield that $\frac{1-\kappa}{2-\kappa}$ is the upper bound of $\eta(\varnothing|\widehat{\theta}^P_{B'})$ as $F \to 0$. At the same time, we have

$$\frac{1-\kappa}{2-\kappa} \leq \underline{\eta} \equiv -\frac{c}{P-c} \Leftrightarrow \kappa \geq \frac{P+c}{P}. \tag{A.106}$$

Consequently, for any $\kappa \geq \frac{P+c}{P}$ we have $\eta(\varnothing|\widehat{\theta}^P_{B'}) < \underline{\eta}$ for any $F > 0$, so that the receiver never invests after $\varnothing$.

Consider the remaining case $\kappa < \frac{P+c}{P}$. Then,

$$\lim_{F\to 0} \eta(\varnothing|\widehat{\theta}^P_{B'}) = \frac{1-\kappa}{2-\kappa} > \underline{\eta}, \tag{A.107}$$

where the equality is by (A.105) and the inequality is by (A.106). Besides, given that $\widehat{\theta}^P_{B'}$ converges to $0.5\bar{\theta}$ as $F \to \bar{\theta}\frac{P-c}{4-\kappa}$ (by (A.43) and the continuity of $\widehat{\theta}^P_{B'}$ in F by Lemma A.13, it holds

$$\lim_{F\to\bar{\theta}\frac{P-c}{4-\kappa}} \eta(\varnothing|\widehat{\theta}^P_{B'}) = \lim_{\widehat{\theta}^P_{B'}\to 0.5\bar{\theta}} \frac{(1-\kappa)(\bar{\theta} - 2\widehat{\theta}^P_{B'})}{(2-\kappa)\bar{\theta} - (4-3\kappa)\widehat{\theta}^P_{B'}} = 0 < \underline{\eta}. \tag{A.108}$$

By (A.100), (A.107), (A.108) and the intermediate value theorem for any $\kappa < \frac{P+c}{P}$ there must exist a threshold value $\widetilde{F}^P \in \left(0, \bar{\theta}\frac{P-c}{4-\kappa}\right)$ such that it should hold $\eta(\varnothing|\widehat{\theta}^P_{B'}) \leq \underline{\eta}$ for any $F \in \left[\widetilde{F}^P, \bar{\theta}\frac{P-c}{4-\kappa}\right]$, and $\eta(\varnothing|\widehat{\theta}^P_{B'}) > \underline{\eta}$ for any $F \in \left(0, \widetilde{F}^P\right)$. Since furthermore $F \geq \bar{\theta}\frac{P-c}{4-\kappa}$ is excluded due to $\widehat{\theta}^P_{N'} < \bar{\theta}$ and Lemma A.8, the incentive constraint $\eta(\varnothing|\widehat{\theta}^P_{B'}) > \underline{\eta}$ under $\kappa < \frac{P+c}{P}$ holds if and only if $F < \widetilde{F}^P$. This together with the previously established fact that $\eta(\varnothing|\widehat{\theta}^P_{B'}) < \underline{\eta}$ for $\kappa \geq \frac{P+c}{P}$ leads to the claim.

(ii) First, note that the receiver never invests after $\varnothing$ (if on the equilibrium path) if $\widehat{\theta}^S_{N'} = \bar{\theta}$, since then only types in state B' refrain in equilibrium so that $\eta(\varnothing) = 0$. Consider the remaining case $\widehat{\theta}^S_{N'} < \bar{\theta}$. By the same argument as in the proof of point (i), we obtain

$$\eta(\varnothing|\widehat{\theta}^{S}_{B'}) = \frac{(1-\kappa)(\bar{\theta}-2\widehat{\theta}^{S}_{B'})}{(2-\kappa)\bar{\theta}-(4-3\kappa)\widehat{\theta}^{S}_{B'}}, \tag{A.109}$$

$$\frac{\partial\eta(\varnothing|\widehat{\theta}^{S}_{B'})}{\partial F} < 0. \tag{A.110}$$

Besides, from Corollary A.2 it follows

$$\lim_{F\to 0}\widehat{\theta}^{S}_{B'} = \lim_{F\to 0}\frac{F}{\eta(m_{N'}|\widehat{\theta}^{S}_{B'})(P-c)} = 0, \tag{A.111}$$

where the last equality holds since $\eta(m_{N'}|\widehat{\theta}^{S}_{B'})$ is constant by (A.77). Then, by the analogous arguments as in the proof of point (i), there must exist a threshold value $\widetilde{F}^{S} \in \left(0, \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}\right)$ such that it should hold $\eta(\varnothing|\widehat{\theta}^{S}_{B'}) > \underline{\eta}$ if and only if $\kappa < \frac{P+c}{P}$ and $F \in \left(0, \widetilde{F}^{S}\right)$. ■

Lemma A.15 *(i) A pooling equilibrium is unresponsive if and only if either of the following holds:*

- $F \geq 0.5\bar{\theta}(P-c)$,
- $\kappa < \frac{P+c}{P}$ *and* $F < \widetilde{F}^{P}$, *where* $\widetilde{F}^{P} \in \left(0, \bar{\theta}\frac{P-c}{4-\kappa}\right)$ *is some threshold value.*

(ii) A separating equilibrium is unresponsive if and only if either of the following holds:

- $F \geq \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}$,
- $\kappa < \frac{P+c}{P}$ *and* $F < \widetilde{F}^{S}$, *where* $\widetilde{F}^{S} \in \left(0, \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}\right)$ *is some threshold value.*

(iii) If an equilibrium is unresponsive, then the receiver obtains an expected payoff of $0.5(P+c)$. *Any responsive equilibrium yields a strictly higher expected payoff.*

Proof. (i),(ii): By Lemma A.5 a pooling or separating equilibrium can be unresponsive only in two cases:

Case 1: $\widehat{\theta}^{P}_{B'} = \bar{\theta}$, i.e. the sender always sends an investment-inducing message. This condition holds if and only if $F \geq 0.5\bar{\theta}(P-c)$ in the pooling equilibrium, and if and only $F \geq \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}$ in the separating equilibrium.

Case 2: $\widehat{\theta}^{P}_{B'} < \bar{\theta}$ while $\eta(\varnothing) > \underline{\eta}$, i.e. the receiver invests conditional on any sender's action in equilibrium. The corresponding necessary and sufficient conditions for this case are given by Lemma A.14.

Combining Cases 1 and 2 together leads to statements (i) and (ii) of the lemma.

(iii) Denote the receiver's payoff in an unresponsive equilibrium as $\underline{U}^r$. Then, by the law of total probability

$$\begin{aligned} E[\underline{U}^r] &= \Pr[I \cap G]P + \Pr[I \cap B]c \\ &= \Pr[G]P + \Pr[B]c \\ &= 0.5(P+c), \end{aligned} \tag{A.112}$$

where the second inequality is by the fact that the receiver always chooses investment.

Next, consider any responsive equilibrium. By Proposition 5 it must be payoff-equivalent to either pooling or separating equilibrium. Hence, without loss of generality assume that the equilibrium is either of these types. Since the equilibrium is assumed to be responsive, the receiver must invest conditional on no refrainment, and abstain conditional on refrainment. Denote the event that the sender chooses to refrain from advice in the considered equilibrium as $\widetilde{\varnothing}$. Then,

$$\begin{aligned} E[U^r] &= E[U^r(I)|\neg\widetilde{\varnothing}] \Pr[\neg\widetilde{\varnothing}] + 0 \cdot \Pr[\widetilde{\varnothing}] \\ &> E[U^r(I)|\neg\widetilde{\varnothing}] \Pr[\neg\widetilde{\varnothing}] + E[U^r(I)|\widetilde{\varnothing}] \Pr[\widetilde{\varnothing}] \\ &= E[U^r(I)] \\ &= 0.5(P+c) \\ &= E[\underline{U}^r], \end{aligned} \tag{A.113}$$

where the first equality is by the law of total probability, the inequality follows from $E[U^r(I)|\widetilde{\varnothing}] < 0$ since the receiver prefers to abstain conditional on the sender refraining in equilibrium, and the last equality is by (A.112).

Thus, all unresponsive equilibria yield for the receiver the same expected payoff of $0.5(P+c)$, while all responsive equilibria have a strictly higher expected payoff. This leads to the claim. ■

Lemma A.16 $\widetilde{F}^P > \widetilde{F}^S$.

Proof. Assume by contradiction that $\widetilde{F}^P \leq \widetilde{F}^S$. Consider some $F \in \left[\widetilde{F}^P, \widetilde{F}^S\right]$ and some $\kappa < \frac{P+c}{P}$ such that both pooling and separating equilibria exist (by Propositions 3 and 4), while $\widetilde{F}^P$ and $\widetilde{F}^S$ are well-defined by Lemma A.15. By Lemma A.15 it holds

$$\widetilde{F}^S < \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}. \tag{A.114}$$

Then, by (A.114) and Lemma A.11,

$$\widehat{\theta}^S_{N'} < \bar{\theta}. \tag{A.115}$$

Consequently, by Lemma 6, we must have

$$\widehat{\theta}^P_{N'} < \widehat{\theta}^S_{N'} < \bar{\theta}. \tag{A.116}$$

In turn, this implies:

$$\eta(\varnothing|\widehat{\theta}^P_{B'}) \leq \underline{\eta} \leq \eta(\varnothing|\widehat{\theta}^S_{B'}), \tag{A.117}$$

where the inequalities follow from $F \in \left[\widetilde{F}^P, \widetilde{F}^S\right]$, $\kappa < \frac{P+c}{P}$ and Lemma A.14 (taking into account that $\eta(\varnothing|\widehat{\theta}^x_{B'}) = \underline{\eta}$ if $F = \widetilde{F}^x$, $x = P, S$, which in turn follows from the continuity of $\eta(\varnothing|\widehat{\theta}^x_{B'})$ in $\widehat{\theta}^x_{B'}$ and hence in F by Lemma A.13).

Next, by (A.116) function $\eta(\varnothing|\widehat{\theta}_{B'})$ is generally given by

$$\eta(\varnothing|\widehat{\theta}_{B'}) = \frac{(1-\kappa)(\overline{\theta} - 2\widehat{\theta}_{B'})}{(2-\kappa)\overline{\theta} - (4-3\kappa)\widehat{\theta}_{B'}}$$

in both equilibria (see (A.99) and (A.109)). Note that the right-hand side is strictly decreasing in $\widehat{\theta}_{B'}$ by (A.102). Consequently, since $\widehat{\theta}^P_{N'} < \widehat{\theta}^S_{N'}$ by (A.116) we obtain $\eta(\varnothing|\widehat{\theta}^S_{B'}) < \eta(\varnothing|\widehat{\theta}^P_{B'})$ which contradicts (A.117). ■

Proof of Proposition 6. Denote the interval of F for which the separating equilibrium exists for given κ as $\Phi(\kappa)$. Note that $\Phi(\kappa)$ is empty if and only if $\kappa \geq \frac{2(P+c)}{2P+c}$ by Proposition 4. Denote further the ex ante receiver's utility in the pooling equilibrium as $U^{r,P}$ and in the separating equilibrium as $U^{r,S}$. Denote $\Delta(F) \equiv U^{r,P}(F) - U^{r,S}(F)$. Note that $\Delta(F)$ is well-defined if and only if $F \in \Phi(\kappa)$ (given that the pooling equilibrium always exists by Proposition 3).

Claim 1. *Assume $\kappa \in \left(0, \frac{2(P+c)}{2P+c}\right)$ so that $\Phi(\kappa)$ is nonempty. There can be only three possible cases:*

- $\Delta(F) \geq 0$ for all $F \in \Phi(\kappa)$;

- $\Delta(F) \leq 0$ for all $F \in \Phi(\kappa)$;

- there exists a threshold $F^* \in \Phi(\kappa)$ such that $\Delta(F) \leq 0$ if $F \in \Phi(\kappa)$ and $F \leq F^*$ (while $\Delta(F) < 0$ for some F), and $\Delta(F) \geq 0$ if $F \in \Phi(\kappa)$ and $F > F^*$ (while $\Delta(F) > 0$ for some F).

Proof. Let us consider properties of $\Delta(F)$ under two parameter cases depending on whether the receiver invests conditional on $\varnothing$.

Case 1. $F < \widetilde{F}^P$ and $\kappa < \frac{P+c}{P}$.

In this case, both equilibria exist while the pooling equilibrium is unresponsive by Lemma A.15(i). Then, by Lemma A.15(iii) the separating equilibrium must yield a (weakly) higher expected payoff, i.e. $\Delta(F) \leq 0$.

Case 2. $\left(\kappa \in \left[\frac{P+c}{P}, \frac{2(P+c)}{2P+c}\right) \wedge F \in \Phi(\kappa)\right)$ or $\left(\kappa < \frac{P+c}{P} \wedge F \geq \widetilde{F}^P\right)$.

Note that in this case both equilibria exist. Besides, note that if $\left(\kappa < \frac{P+c}{P} \wedge F \geq \widetilde{F}^P\right)$, then $F > \widetilde{F}^S$ by Lemma A.16. Consequently, by Lemma A.14 in Case 2 the receiver must abstain conditional on $\varnothing$ in both equilibria (if on the equilibrium path). Hence, the ex ante receiver's utility given any equilibrium cutoffs $\widehat{\theta}_{B'}$ and $\widehat{\theta}_{N'}$ is

$$\begin{aligned}
E[U^r] &= \Pr[I \cap G]P + \Pr[I \cap B]c \\
&= \left(\sum_{i^s \in \{G',N',B'\}} \Pr[I|i^s \cap G] \Pr[i^s|G] \right) \Pr[G] \cdot P \\
&\quad + \left(\sum_{i^s \in \{G',N',B'\}} \Pr[I|i^s \cap B] \Pr[i^s|B] \right) \Pr[B] \cdot c \\
&= \left(\sum_{i^s \in \{G',N',B'\}} \Pr[I|i^s] \Pr[i^s|G] \right) \Pr[G] \cdot P \\
&\quad + \left(\sum_{i^s \in \{G',N',B'\}} \Pr[I|i^s] \Pr[i^s|B] \right) \Pr[B] \cdot c \\
&= \kappa(0.5P + 0.5c\frac{\widehat{\theta}_{B'}}{\bar{\theta}}) + (1-\kappa)\frac{\widehat{\theta}_{N'}}{\bar{\theta}}(0.5P + 0.5c), \qquad \text{(A.118)}
\end{aligned}$$

where the second equality is by the law of total probability, the third equality is by the fact that sender messages in equilibrium (and hence the receiver's actions) do not depend on the true state of the world once his information state i^s is conditioned upon, and the fourth equality is obtained by substituting $\Pr[I|G'] = 1$, $\Pr[I|N'] = \frac{\widehat{\theta}_{N'}}{\bar{\theta}}$, $\Pr[I|B'] = \frac{\widehat{\theta}_{B'}}{\bar{\theta}}$, $\Pr[G'|G] = \Pr[B'|B] = \kappa$, $\Pr[G'|B] = \Pr[B'|G] = 0$, and $\Pr[N'|G] = \Pr[N'|B] = 1 - \kappa$.

By Lemma 6 we have the following possible equilibrium cases: 1) $\widehat{\theta}^S_{N'} < \bar{\theta}$ and $\widehat{\theta}^P_{N'} < \bar{\theta}$; 2) $\widehat{\theta}^S_{N'} = \bar{\theta}$ and $\widehat{\theta}^P_{N'} < \bar{\theta}$; 3) $\widehat{\theta}^S_{N'} = \bar{\theta}$ and $\widehat{\theta}^P_{N'} = \bar{\theta}$. Let us consider these cases sequentially.

Case 2.1: $\widehat{\theta}^S_{N'} < \bar{\theta}$ and $\widehat{\theta}^P_{N'} < \bar{\theta}$.

By (A.118) and $\widehat{\theta}_{N'} = 2\widehat{\theta}_{B'}$ (by Lemmas A.6 and A.9) we obtain

$$\begin{aligned}
\Delta(F) &= U^{r,P} - U^{r,S} = \kappa(0.5P + 0.5c\frac{\widehat{\theta}^P_{B'}}{\bar{\theta}}) + (1-\kappa)\frac{2\widehat{\theta}^P_{B'}}{\bar{\theta}}(0.5P + 0.5c) \\
&\quad -\kappa(0.5P + 0.5c\frac{\widehat{\theta}^S_{B'}}{\bar{\theta}}) - (1-\kappa)\frac{2\widehat{\theta}^S_{B'}}{\bar{\theta}}(0.5P + 0.5c) \\
&= \kappa\frac{0.5c}{\bar{\theta}}(\widehat{\theta}^P_{B'} - \widehat{\theta}^S_{B'}) + (1-\kappa)\frac{P+c}{\bar{\theta}}(\widehat{\theta}^P_{B'} - \widehat{\theta}^S_{B'}) \\
&= (\widehat{\theta}^P_{B'} - \widehat{\theta}^S_{B'})(\frac{0.5\kappa c + (1-\kappa)(P+c)}{\bar{\theta}}). \qquad \text{(A.119)}
\end{aligned}$$

By Lemma 6 the first term in the RHS is strictly negative. At the same time, the second term is strictly positive given that $\kappa < \frac{2(P+c)}{2P+c}$ by assumption. Consequently, in Case 2.1

$$\Delta(F) < 0. \qquad \text{(A.120)}$$

Case 2.2: $\widehat{\theta}^S_{N'} = \bar{\theta}$ and $\widehat{\theta}^P_{N'} < \bar{\theta}$.

For this case, let us just show that Δ strictly increases in F. Indeed, by (A.118), given that $\widehat{\theta}^P_{N'} = 2\widehat{\theta}^P_{B'}$ and $\widehat{\theta}^S_{N'} = \bar{\theta}$, we have

$$\begin{aligned}
\Delta(F) &= U^{r,P} - U^{r,S} = \kappa(0.5P + 0.5c\frac{\widehat{\theta}^P_{B'}}{\bar{\theta}}) + (1-\kappa)\frac{2\widehat{\theta}^P_{B'}}{\bar{\theta}}(0.5P + 0.5c) \\
&\quad -\kappa(0.5P + 0.5c\frac{\widehat{\theta}^S_{B'}}{\bar{\theta}}) - (1-\kappa)\frac{1}{\bar{\theta}}(0.5P + 0.5c) \\
&= -0.5\kappa c\frac{\widehat{\theta}^S_{B'}}{\bar{\theta}} + (P(1-\kappa) + c(1-0.5\kappa))\frac{\widehat{\theta}^P_{B'}}{\bar{\theta}} \\
&\quad -(1-\kappa)(0.5P + 0.5c). \qquad \text{(A.121)}
\end{aligned}$$

Then, where the derivative exists,

$$\frac{\partial \Delta}{\partial F} = \frac{-0.5\kappa c}{\bar{\theta}}\frac{\partial \widehat{\theta}^S_{B'}}{\partial F} + (P(1-\kappa) + c(1-0.5\kappa))\frac{1}{\bar{\theta}}\frac{\partial \widehat{\theta}^P_{B'}}{\partial F}. \qquad \text{(A.122)}$$

By (A.103) and (A.65) we have that $\frac{\partial \widehat{\theta}^S_{B'}}{\partial F}$ and $\frac{\partial \widehat{\theta}^P_{B'}}{\partial F}$ are strictly positive (where these derivatives exist). Besides, the term $P(1-\kappa) + c(1-0.5\kappa)$ is strictly positive since $\kappa < \frac{2(P+c)}{2P+c}$ by assumption. Hence, by (A.122), where $\frac{\partial \Delta}{\partial F}$ exists, it is strictly positive. Finally, since $\Delta(F)$ is continuous in both cutoffs and hence in F by Lemma A.13, while being also differentiable in F at a given cutoff unless this cutoff takes values of $0.5\bar{\theta}$ or $\bar{\theta}$ (see the proof of Lemma A.13), we obtain that in Case 2.2 Δ strictly increases in F.

Case 2.3: $\widehat{\theta}^S_{N'} = \bar{\theta}$ and $\widehat{\theta}^P_{N'} = \bar{\theta}$.

By (A.118) we obtain

$$\begin{aligned}
\Delta(F) &= U^{r,P} - U^{r,S} = \kappa(0.5P + 0.5c\frac{\widehat{\theta}^P_{B'}}{\bar{\theta}}) - \kappa(0.5P + 0.5c\frac{\widehat{\theta}^S_{B'}}{\bar{\theta}}) \\
&= \kappa\frac{0.5c}{\bar{\theta}}(\widehat{\theta}^P_{B'} - \widehat{\theta}^S_{B'}) \geq 0, \qquad \text{(A.123)}
\end{aligned}$$

where the last inequality is by Lemma 6.

Finally, let us combine Case 1, and Cases 2.1, 2.2 and 2.3 together. Let us denote the intervals of F corresponding to these cases for given κ as $\Phi_1(\kappa)$, $\Phi_{21}(\kappa)$, $\Phi_{22}(\kappa)$ and $\Phi_{23}(\kappa)$, respectively. Since by Proposition 4 the separating equilibrium exists if and only if either of the conditions for Case 1 or for Case 2 is satisfied (and the pooling equilibrium always exists by Proposition 3), these intervals cover the range of all possible values of F where both pooling and separating equilibria simultaneously exist for given κ, i.e. $\Phi(\kappa) = \Phi_1(\kappa) \cup \Phi_{21}(\kappa) \cup \Phi_{22}(\kappa) \cup \Phi_{23}(\kappa)$. Note also that since $\widehat{\theta}_{B'}$ and hence $\widehat{\theta}_{N'}$ are continuously increasing in F in either equilibrium by (A.103), (A.65) and Lemma A.13, we must have that $\Phi_{21}(\kappa)$ (if nonempty) lies to the left of $\Phi_{22}(\kappa)$ (if nonempty), which in turn lies to the left of $\Phi_{23}(\kappa)$ (if nonempty). Besides, by construction, Φ_1 (if nonempty) must lie to the left of any of the other intervals. At the same time, from the above analysis of the cases we have

$$\begin{aligned} \forall F &\in \Phi_1 : \Delta(F) \leq 0, && \text{(A.124)} \\ \forall F &\in \Phi_{21} : \Delta(F) < 0, && \text{(A.125)} \\ \forall F &\in \Phi_{22} : \Delta \text{ strictly increases in } F, && \text{(A.126)} \\ \forall F &\in \Phi_{23} : \Delta(F) \geq 0 && \text{(A.127)} \end{aligned}$$

This together with the disposition of the intervals described above leads to Claim 1.

Claim 2. *Assume* $\kappa \in \left(0, \frac{2(P+c)}{2P+c}\right)$ *so that* $\Phi(\kappa)$ *is nonempty. Then, there exists a value of* $F > 0$ *within* $\Phi(\kappa)$ *such that* $\Delta(F) < 0$.

Proof. In what follows, we refer to Cases 1 and 2.1-2.3 from the proof of Claim 1.

Consider first $\kappa < \frac{P+c}{P}$. Then, by Propositions 3 and 4, both the pooling and the separating equilibrium exist for any F. Consider any $F' \in \left(\widetilde{F}^S, \widetilde{F}^P\right)$, which should exist by Lemma A.16. Moreover, set F to be sufficiently close to $\widetilde{F}^S$ so that

$$F < \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}$$

(which is feasible by Lemma A.15). Consequently, by Lemma A.11

$$\widehat{\theta}^S_{B'} < \bar{\theta}. \tag{A.128}$$

Then, since $F' < \widetilde{F}^P$, the pooling equilibrium is unresponsive (by Lemma A.15). At the same time, it follows from $F' > \widetilde{F}^S$ and Lemma A.14 that the receiver should abstain conditional on refrainment in the separating equilibrium (if on the equilibrium path), which together with (A.128) implies that the separating equilibrium is responsive. Then, $\Delta(F) < 0$ by Lemma A.15(iii).

Consider the remaining case $\kappa \in \left[\frac{P+c}{P}, \frac{2(P+c)}{2P+c}\right)$. Then, by Proposition 4 the separating equilibrium exists if and only if $F < \bar{\theta}\,(P+c)\,\frac{(1-\kappa)}{\kappa}$. Moreover, by Lemmas A.8 and Lemma A.11 the conditions for Case 2.1 in the proof of Claim 1 (including the general parameter restrictions for Case 2) will then be satisfied for sufficiently small $F \in \left(0, \bar{\theta}\,(P+c)\,\frac{(1-\kappa)}{\kappa}\right)$. For such values of F it will hold $\Delta(F) < 0$ by (A.125).

Claim 3. *Assume* $\kappa \in \left(0, \frac{2(P+c)}{2P+c}\right)$ *so that* $\Phi(\kappa)$ *is nonempty. Then, there exists a value of* $F > 0$ *within* $\Phi(\kappa)$ *such that* $\Delta(F) > 0$.

Proof. Consider first $\kappa < \frac{P+c}{P}$. Then, by Propositions 3 and 4, both the pooling and the separating equilibrium exist for any $F > 0$. Take $F \in \left[\max\left\{\frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}, \widetilde{F}^P\right\}, 0.5\bar{\theta}(P-c)\right)$. Then, by Lemmas A.8, A.11 and A.14 we have $\widehat{\theta}^S_{B'} = \widehat{\theta}^S_{N'} = \bar{\theta}$, i.e. the separating equilibrium is unresponsive, while $\widehat{\theta}^P_{B'} < \widehat{\theta}^P_{N'} \leq \bar{\theta}$ and the receiver abstains after $\varnothing$ in the pooling equilibrium (i.e. the pooling equilibrium is responsive). Then, $\Delta(F) > 0$ by Lemma A.15(iii).

Consider the remaining case $\kappa \in \left[\frac{P+c}{P}, \frac{2(P+c)}{2P+c}\right)$. Then, by Proposition 4 the separating equilibrium exists if and only if $F < \bar{\theta}\,(P+c)\,\frac{(1-\kappa)}{\kappa}$. At the same time, if F goes to $\bar{\theta}\,(P+c)\,\frac{(1-\kappa)}{\kappa}$, then $\eta(\widetilde{m})$ converges to $\underline{\eta}$ (by (A.75) and the fact that $\eta(\widetilde{m})$ is continuous in $\widehat{\theta}^S_{B'}$ and hence in F by Lemma A.13). Consequently, the receiver's utility conditional on $\widetilde{m}$ is going to 0. Then, in the limit the only ex ante (strictly) profitable message in the separating equilibrium is $\bar{m}$ (sent only by types in state G'), that is

$$\lim_{F\to\bar{\theta}(P+c)\frac{(1-\kappa)}{\kappa}} U^{r,S} = E[U^{r,S}|\bar{m}]\Pr[\bar{m}] = 0.5P\kappa. \tag{A.129}$$

It follows, denoting $\overrightarrow{\theta}^P = \lim_{F\to\bar{\theta}(P+c)\frac{(1-\kappa)}{\kappa}} \widehat{\theta}^P_{B'}$,

$$\begin{aligned}
&\lim_{F\to\bar{\theta}(P+c)\frac{(1-\kappa)}{\kappa}} \Delta(F) \\
&= \kappa(0.5P + 0.5c\frac{\overrightarrow{\theta}^P}{\bar{\theta}}) + (1-\kappa)\frac{2\overrightarrow{\theta}^P}{\bar{\theta}}(0.5P+0.5c) - 0.5P\kappa \\
&= \frac{\overrightarrow{\theta}^P}{\bar{\theta}}(P(1-\kappa) + c(1-0.5\kappa)) > 0,
\end{aligned} \tag{A.130}$$

where the inequality follows since $\kappa < \frac{2(P+c)}{2P+c}$ by assumption and $\overrightarrow{\theta}^P \neq 0$ (since by Corollary A.1 $\widehat{\theta}^P_{B'} = \frac{F}{\eta(\bar{m})(P-c)}$ which is bounded from 0).

Thus, there exists $F > 0$ within $\Phi(\kappa)$ such that $\Delta(F) > 0$.

Claim 4. *Assume $\kappa \in \left(0, \frac{2(P+c)}{2P+c}\right)$ so that $\Phi(\kappa)$ is nonempty. Then, there exists $F^* \in \Phi(\kappa)$ such that $\Delta(F) \leq 0$ if $F \in \Phi(\kappa)$ and $F \leq F^*$ (while $\Delta(F) < 0$ for some F), and $\Delta(F) \geq 0$ if $F \in \Phi(\kappa)$ and $F > F^*$ (while $\Delta(F) > 0$ for some F).*

Proof. The claim follows jointly from Claims 1-3. In particular, Claim 2 rules out the first case listed in Claim 1, while Claim 3 rules out the second case. Hence, the only possible case is the third case listed in Claim 1. ■

Proof of Proposition 7. By Lemma 6, the cutoffs in the pooling equilibrium in each state are (weakly) lower than in the separating equilibrium. Hence, for a given sender type θ' in a given state $i^s \in \{B', N'\}$ there can be only three possible cases, as listed below.

Case 1. $\theta' \leq \widehat{\theta}^P_{i^s} \leq \widehat{\theta}^S_{i^s}$. Then, by Lemma 4, θ' would send an investment-inducing message in both equilibria. He would strictly prefer the separating equilibrium due to a strictly lower expected guilt, since $\eta(\widetilde{m}) < 0.5 \leq \eta(\bar{m})$ by (A.44) and (A.80).

Case 2. $\widehat{\theta}^P_{i^s} < \theta' \leq \widehat{\theta}^S_{i^s}$. Then, by Lemma 4, θ' would send an investment-inducing message in the separating equilibrium (leading to a positive utility) while would obtain a utility of 0 in the pooling equilibrium. Hence, he would again prefer the separating equilibrium, with a strict preference if $\theta' < \widehat{\theta}^S_{i^s}$.

Case 3. $\widehat{\theta}_{i^s}^P \leq \widehat{\theta}_{i^s}^S < \theta'$. Then, by Lemma 4, θ' would get a utility of 0 under both equilibria.

Finally, all types in state G' would be indifferent between the pooling and the separating equilibrium, since they always obtain F by Lemma 3.

Thus, the expected utility of each sender's type is always (weakly) higher in the separating equilibrium, and strictly so at least for some types in states B' and N'. ■

A.8 Welfare effects of ex ante information quality

Lemma A.17 *In any equilibrium,* $E[U^r] = \max\{\psi, 0.5(P+c)\}$*, where*

$$\psi = \kappa(0.5P + 0.5c\frac{\widehat{\theta}_{B'}}{\overline{\theta}}) + (1-\kappa)\frac{\widehat{\theta}_{N'}}{\overline{\theta}}0.5(P+c). \tag{A.131}$$

Proof. Consider three possible cases depending on the value of the cutoffs and $\Pr[G|\varnothing]$.

Case 1: $\widehat{\theta}_{B'} = \widehat{\theta}_{N'} = \bar{\theta}$. Then, the receiver always invests independently of the state so that

$$E[U^r] = 0.5(P+c) = \psi_{|\widehat{\theta}_{B'}=\widehat{\theta}_{N'}=\bar{\theta}},$$

which also implies $E[U^r] = \max\{\psi, 0.5(P+c)\}$.

Case 2: $\widehat{\theta}_{B'} < \widehat{\theta}_{N'} \leq \bar{\theta}$ and $\Pr[G|\varnothing] \leq \underline{\eta}$. Then, the receiver abstains conditional on $\varnothing$, i.e. the equilibrium is responsive. Consequently,

$$E[U^r] = \psi > 0.5(P+c)$$

where the equality is by (A.118), and the inequality is by Lemma A.15(iii). Hence, we again have $E[U^r] = \max\{\psi, 0.5(P+c)\}$.

Case 3: $\widehat{\theta}_{B'} < \widehat{\theta}_{N'} \leq \bar{\theta}$ and $\Pr[G|\varnothing] > \underline{\eta}$. Then, the receiver invests conditional on $\varnothing$, i.e. the equilibrium is unresponsive so that by Lemma A.15(iii)

$$E[U^r] = 0.5(P+c). \tag{A.132}$$

Note that $\Pr[G|\varnothing] > \underline{\eta}$ implies $\widehat{\theta}_{N'} < \bar{\theta}$ since otherwise $\Pr[G|\varnothing] = 0$. Then, by (A.99) we have

$$\Pr[G|\varnothing] = \frac{(1-\kappa)(\overline{\theta} - 2\widehat{\theta}_{B'}^P)}{(2-\kappa)\overline{\theta} - (4-3\kappa)\widehat{\theta}_{B'}^P}$$

so that

$$\begin{aligned} \Pr[G|\varnothing] &> \underline{\eta} \\ &\Leftrightarrow \frac{(1-\kappa)(\overline{\theta} - 2\widehat{\theta}_{B'}^P)}{(2-\kappa)\overline{\theta} - (4-3\kappa)\widehat{\theta}_{B'}^P} > \frac{-c}{P-c} \\ &\Leftrightarrow \psi < 0.5(P+c). \end{aligned} \tag{A.133}$$

Then, (A.132) and (A.133) imply $E[U^r] = \max\{\psi, 0.5(P+c)\}$

Thus, we have shown that $E[U^r] = \max\{\psi, 0.5(P+c)\}$ in all possible cases. ■

Corollary A.4 *In both pooling and separating equilibria, $E[U^r]$ and ψ are continuous in both $F \in (0,\infty)$ and $\kappa \in (0,1)$.*

Proof. By (A.131) ψ is continuous in κ (as a direct argument) and in the cutoffs $\widehat{\theta}_{B'}(F,\kappa)$ and $\widehat{\theta}_{N'}(F,\kappa)$. At the same time, $\widehat{\theta}_{B'}(F,\kappa)$ and $\widehat{\theta}_{N'}(F,\kappa)$ are continuous in both F and κ by Lemma A.13 and Corollary A.3, hence ψ is also continuous in both F and κ. Finally, since by Lemma A.17 $E[U^r]$ is continuous in ψ, it is also continuous in both F and κ . ■

Proof of Proposition 8.

(i) *The pooling equilibrium.*

In what follows, we consider function ψ defined in Lemma A.17:

$$\psi = \kappa(0.5P + 0.5c\frac{\widehat{\theta}^P_{B'}(\kappa)}{\bar{\theta}}) + (1-\kappa)\frac{\widehat{\theta}^P_{N'}(\kappa)}{\bar{\theta}}(0.5P + 0.5c). \tag{A.134}$$

Claim 1. *In the pooling equilibrium, there exists threshold $\kappa' \in \left(\max\left\{\frac{4F-\bar{\theta}(P-c)}{F}, 0\right\}, 1\right)$ such that ψ is strictly decreasing in κ on some interval K if and only if for any $\kappa \in K$ it holds*:

- $\frac{\bar{\theta}(P-c)^2}{2(3P+c)} < F < \frac{1}{6}\bar{\theta}(-2c + P + \sqrt{P^2 + 2c(P-c)})$,
- $\kappa \in \left(\max\left\{\frac{4F-\bar{\theta}(P-c)}{F}, 0\right\}, \kappa'\right)$.

Let us consider the following possible cases depending on the possible parameter values.

Case 1: $F \geq 0.5\bar{\theta}(P-c)$ so that $\widehat{\theta}^P_{N'} = \bar{\theta}$ and $\widehat{\theta}^P_{B'} = \bar{\theta}$ (by Lemma A.8). Then, by (A.134)

$$\begin{aligned} \psi &= \kappa(0.5P + 0.5c) + (1-\kappa)(0.5P + 0.5c) \\ &= 0.5P + 0.5c. \end{aligned} \tag{A.135}$$

Hence, in this case

$$\frac{d\psi}{d\kappa} = 0. \tag{A.136}$$

Case 2: $F < 0.5\bar{\theta}(P-c)$ and $\left(F > \bar{\theta}\frac{P-c}{4-\kappa} \iff \kappa \in \left(0, \min\left\{\frac{4F-\bar{\theta}(P-c)}{F}, 1\right\}\right)\right)$. Then, $\widehat{\theta}^P_{N'} = \bar{\theta}$ and $\widehat{\theta}^P_{B'} \in (0.5, \bar{\theta})$ by Lemma A.8.

In this case, by (A.134)

$$\psi = \kappa(0.5P + 0.5c\frac{\widehat{\theta}^P_{B'}}{\bar{\theta}}) + (1-\kappa)(0.5P + 0.5c). \tag{A.137}$$

We have

$$\frac{d\psi}{d\kappa} = \frac{\partial\psi}{\partial\kappa} + \frac{\partial\psi}{\partial\widehat{\theta}^P_{B'}}\frac{\partial\widehat{\theta}^P_{B'}}{\partial\kappa}, \tag{A.138}$$

where

$$\frac{\partial\psi}{\partial\kappa} = \frac{-c(\bar{\theta}-\widehat{\theta}^P_{B'})}{2\bar{\theta}} > 0, \tag{A.139}$$

$$\frac{\partial\psi}{\partial\widehat{\theta}^P_{B'}} = \frac{c\kappa}{2\bar{\theta}} < 0. \tag{A.140}$$

At the same time,

$$\frac{\partial\widehat{\theta}^P_{B'}}{\partial\kappa} = -\frac{\partial\varpi(\widehat{\theta}^P_{B'})/\partial\kappa}{\partial\varpi(\widehat{\theta}^P_{B'})/\partial\widehat{\theta}^P_{B'}} = -\frac{(\bar{\theta}-\widehat{\theta}^P_{B'})\widehat{\theta}^P_{B'}}{(2-\kappa)\bar{\theta}} < 0, \tag{A.141}$$

where the first equality is by the implicit function theorem, and the second equality is obtained by substituting for $\eta(\bar{m}|\widehat{\theta}^P_{B'})$ in $\varpi(\widehat{\theta}^P_{B'})$ (see (A.33)) and simplifying. Finally, (A.138)-(A.141) imply that in Case 2

$$\frac{d\psi}{d\kappa} > 0. \tag{A.142}$$

Case 3: $F < \bar{\theta}\frac{P-c}{4-\kappa} \iff \kappa \in \left(\max\left\{\frac{4F-\bar{\theta}(P-c)}{F}, 0\right\}, 1\right)$. Then, $\bar{\theta} > \widehat{\theta}^P_{N'} = 2\widehat{\theta}^P_{B'}$ by Lemma A.8.

Step 1. Let us show that

$$\frac{d^2\psi}{d\kappa^2} > 0. \tag{A.143}$$

Denoting $\left(\widehat{\theta}^P_{B'}\right)' = \frac{\partial\widehat{\theta}^P_{B'}}{\partial\kappa}$ and $\left(\widehat{\theta}^P_{B'}\right)'' = \frac{\partial^2\widehat{\theta}^P_{B'}}{\partial\kappa^2}$ we have

$$\begin{aligned}\frac{d\psi}{d\kappa} &= \frac{d\left(\kappa(0.5P+0.5c\frac{\widehat{\theta}^P_{B'}}{\bar{\theta}}) + (1-\kappa)\frac{2\widehat{\theta}^P_{B'}}{\bar{\theta}}(0.5P+0.5c)\right)}{d\kappa} \\ &= \frac{P\bar{\theta} - (2P+c)\widehat{\theta}^P_{B'} + (2(P+c)-(2P+c)\kappa)\left(\widehat{\theta}^P_{B'}\right)'}{2\bar{\theta}},\end{aligned} \tag{A.144}$$

so that

$$\frac{d^2\psi}{d\kappa^2} = \frac{-2(2P+c)\left(\widehat{\theta}^P_{B'}\right)' + (2(P+c)-(2P+c)\kappa)\left(\widehat{\theta}^P_{B'}\right)''}{2\bar{\theta}}. \tag{A.145}$$

Consider $\left(\widehat{\theta}^P_{B'}\right)'$ and $\left(\widehat{\theta}^P_{B'}\right)''$. By Lemma A.6, $\varpi(\widehat{\theta}^P_{B'}) = 0$. Consequently, by the implicit function theorem

$$\left(\widehat{\theta}^P_{B'}\right)' = -\frac{\partial\varpi(\widehat{\theta}^P_{B'})/\partial\kappa}{\partial\varpi(\widehat{\theta}^P_{B'})/\partial\widehat{\theta}^P_{B'}}. \tag{A.146}$$

Substituting and simplifying we obtain

$$\left(\widehat{\theta}_{B'}^{P}\right)' = -\frac{2\left(\widehat{\theta}_{B'}^{P}\right)^2(\bar{\theta}-\widehat{\theta}_{B'}^{P})}{\kappa^2\bar{\theta}^2+4(1-\kappa)\kappa\bar{\theta}\widehat{\theta}_{B'}^{P}+2(1-\kappa)(4-3\kappa)\left(\widehat{\theta}_{B'}^{P}\right)^2} < 0. \tag{A.147}$$

Further, by the chain rule

$$\left(\widehat{\theta}_{B'}^{P}\right)'' = \frac{\partial\left(\widehat{\theta}_{B'}^{P}\right)'}{\partial\kappa} + \frac{\partial\left(\widehat{\theta}_{B'}^{P}\right)'}{\partial\widehat{\theta}_{B'}^{P}}\left(\widehat{\theta}_{B'}^{P}\right)'. \tag{A.148}$$

Calculating this expression by (A.147) and substituting it together with (A.147) into (A.145) we get

$$\frac{d^2\psi}{d\kappa^2} = \frac{4(\bar{\theta}-\widehat{\theta}_{B'}^{P})(\widehat{\theta}_{B'}^{P})^2}{\bar{\theta}(\kappa^2\bar{\theta}^2+4(1-\kappa)\kappa\bar{\theta}\widehat{\theta}_{B'}^{P}+2(4-7\kappa+3\kappa^2)(\widehat{\theta}_{B'}^{P})^2)^3}Z, \tag{A.149}$$

where

$$\begin{aligned} Z &= (P+c)\kappa^3\bar{\theta}^4+8(P+c)(1-\kappa)\kappa^2\bar{\theta}^3\widehat{\theta}_{B'}^{P}-2c(4-3\kappa)^2(1-\kappa)(\widehat{\theta}_{B'}^{P})^4 \\ &\quad +2\kappa\bar{\theta}^2(\widehat{\theta}_{B'}^{P})^2(10P(1-\kappa)^2+c(10-23\kappa+12\kappa^2)) \\ &\quad +8(1-\kappa)(2P(1-\kappa)^2+c(2-7\kappa+4\kappa^2))\bar{\theta}(\widehat{\theta}_{B'}^{P})^3. \end{aligned}$$

The multiple of Z in (A.149) is clearly positive, hence to show the claim we need to prove that $Z > 0$. The first three terms of Z are clearly positive. For the fourth term we have (given that $P > -c$ by Assumption 1)

$$\begin{aligned} 10P(1-\kappa)^2+c(10-23\kappa+12\kappa^2) &> -10c(1-\kappa)^2+c(10-23\kappa+12\kappa^2) \\ &= -c\kappa\,(3-2\kappa) > 0. \end{aligned} \tag{A.150}$$

Finally, for the fifth term we have

$$\begin{aligned} 2P(1-\kappa)^2+c(2-7\kappa+4\kappa^2) &> -2c(1-\kappa)^2+c(2-7\kappa+4\kappa^2) \\ &= -c\kappa\,(3-2\kappa) > 0. \end{aligned} \tag{A.151}$$

Consequently, $Z > 0$ so that by (A.149)

$$\frac{d^2\psi}{d\kappa^2} > 0. \tag{A.152}$$

Step 2. Denote by $\underline{\kappa}$ the infimum of the set of κ corresponding to Case 3:

$$\underline{\kappa} = \max\left\{0, \frac{4F - \bar{\theta}(P-c)}{F}\right\}. \tag{A.153}$$

Let us show that in Case 3

$$\begin{aligned} \lim_{\kappa \to \underline{\kappa}} \frac{d\psi}{d\kappa} &< 0 \Leftrightarrow \\ \frac{\bar{\theta}(P-c)^2}{2(3P+c)} &< F < \frac{1}{6}\bar{\theta}(-2c + P + \sqrt{P^2 + 2c(P-c)}). \end{aligned} \tag{A.154}$$

Consider first the case of $\underline{\kappa} = 0$ which, by (A.153), occurs if and only if $F \leq 0.25\bar{\theta}(P-c)$. By Corollary A.1 and $\widehat{\theta}^P_{B'} < \bar{\theta}$ by assumption,

$$\widehat{\theta}^P_{B'} = \frac{F}{\eta(\bar{m})(P-c)} \tag{A.155}$$

so that

$$\lim_{\kappa \to \underline{\kappa}} \widehat{\theta}^P_{B'} = \frac{F}{\lim_{\kappa \to 0} \eta(\bar{m})(P-c)}. \tag{A.156}$$

At the same time, since $\widehat{\theta}^P_{N'} = 2\widehat{\theta}^P_{B'}$ in the considered case, by (A.29)

$$\eta(\bar{m}|\widehat{\theta}^P_{B'}) = \frac{2\widehat{\theta}^P_{B'}(1-\kappa) + \kappa\bar{\theta}}{\widehat{\theta}^P_{B'}(4-3\kappa) + \kappa\bar{\theta}}. \tag{A.157}$$

Since $\widehat{\theta}^P_{B'}$ is bounded from 0 by (A.155) (and hence the limit of the denominator is not equal to 0), by (A.157) we obtain

$$\lim_{\kappa \to 0} \eta(\bar{m}|\widehat{\theta}^P_{B'}) = \frac{2\lim_{\kappa \to 0} \widehat{\theta}^P_{B'}}{4\lim_{\kappa \to 0} \widehat{\theta}^P_{B'}} = 0.5. \tag{A.158}$$

This together with (A.156) implies

$$\lim_{\kappa \to \underline{\kappa}} \widehat{\theta}^P_{B'} = \frac{2F}{P-c}. \tag{A.159}$$

Substituting this into (A.147) we get (given that the denominator is again bounded from 0)

$$\lim_{\kappa \to \underline{\kappa}} \left(\widehat{\theta}^P_{B'}\right)' = \frac{F}{2(P-c)} - \frac{\bar{\theta}}{4}. \tag{A.160}$$

Finally, from (A.159), (A.160) and (A.144) we obtain

$$\begin{aligned}\lim_{\kappa\to\underline{\kappa}}\frac{d\psi}{d\kappa} &= \frac{P\bar{\theta}-(2P+c)\frac{2F}{P-c}+2(P+c)(\frac{F}{2(P-c)}-\frac{\bar{\theta}}{4})}{2\bar{\theta}}\\ &= \frac{\bar{\theta}(P-c)^2-2F(3P+c)}{4\bar{\theta}(P-c)}. \qquad (A.161)\end{aligned}$$

It follows that $\lim_{\kappa\to\underline{\kappa}}\frac{d\psi}{d\kappa}<0$ if and only if

$$F>\frac{\bar{\theta}(P-c)^2}{2(3P+c)}.$$

Given the initial assumption $\underline{\kappa}=0 \Leftrightarrow F\le 0.25\bar{\theta}(P-c)$, this happens if and only if $F\in\left(\frac{\bar{\theta}(P-c)^2}{2(3P+c)},0.25\bar{\theta}(P-c)\right)$, which is nonempty if $P>-3c$.

Consider the remaining case $\underline{\kappa}\neq 0$, i.e. $\underline{\kappa}=\frac{4F-\bar{\theta}(P-c)}{F}>0$ (see (A.153)). Note that

$$\widehat{\theta}^P_{B'}(\underline{\kappa})=\widehat{\theta}^P_{B'|\underline{\kappa}=\frac{4F-\bar{\theta}(P-c)}{F}}=\widehat{\theta}^P_{B'|F=\bar{\theta}\frac{P-c}{4-\underline{\kappa}}}=0.5\bar{\theta},$$

where the last equality follows from (A.43) and the equilibrium condition $\varpi(\widehat{\theta}^P_{B'})=0$. Consequently, since $\widehat{\theta}^P_{B'}(\kappa)$ is continuous in κ by Lemma A.13, it follows

$$\lim_{\kappa\to\underline{\kappa}}\widehat{\theta}^P_{B'}(\underline{\kappa})=\widehat{\theta}^P_{B'}(\underline{\kappa})=0.5\bar{\theta}. \qquad (A.162)$$

Substituting this into (A.147) we obtain

$$\lim_{\kappa\to\underline{\kappa}}\left(\widehat{\theta}^P_{B'}\right)'=\frac{F^2\bar{\theta}}{2(5F\bar{\theta}(P-c)-(P-c)^2\bar{\theta}^2-8F^2)} \qquad (A.163)$$

Finally, substituting (A.162) and (A.163) into (A.144) yields

$$\lim_{\kappa\to\underline{\kappa}}\frac{d\psi}{d\kappa}=-\frac{(P-c)(6F^2-c(P-c)\bar{\theta}^2-2F\bar{\theta}(P-2c))}{4(5F\bar{\theta}(P-c)-(P-c)^2\bar{\theta}^2-8F^2)}.$$

One can show that this expression is strictly negative in the considered case (i.e. under the restriction $F>0.25\bar{\theta}(P-c)$) if and only if

$$F\in\left(0.25\bar{\theta}(P-c),\frac{1}{6}(-2c\bar{\theta}+P\bar{\theta}+\bar{\theta}\sqrt{P^2+2cP-2c^2})\right)$$

This interval is nonempty if and only if $P>-3c$.

Combining the cases $\underline{\kappa} = 0$ and $\underline{\kappa} \neq 0$ together, we get that in Case 3

$$\begin{aligned} \lim_{\kappa \to \underline{\kappa}} \frac{d\psi}{d\kappa} &< 0 \Leftrightarrow \\ F &\in \left(\frac{\bar{\theta}(P-c)^2}{2(3P+c)}, \frac{1}{6}(-2c\bar{\theta} + P\bar{\theta} + \bar{\theta}\sqrt{P^2+2cP-2c^2}) \right), \end{aligned} \tag{A.164}$$

with the interval being nonempty if and only if $P > -3c$.

Step 3. Let us show that in Case 3

$$\lim_{\kappa \to 1} \frac{d\psi}{d\kappa} > 0. \tag{A.165}$$

From (A.147) we have

$$\lim_{\kappa \to 1} \left(\widehat{\theta}^P_{B'} \right)' = -\frac{2(\bar{\theta} - \widehat{\theta}^P_{B'})(\widehat{\theta}^P_{B'})^2}{\bar{\theta}^2}. \tag{A.166}$$

Equations (A.144) and (A.166) lead to

$$\lim_{\kappa \to 1} \frac{d\psi}{d\kappa} = \frac{P\bar{\theta}^2(\bar{\theta} - 2\widehat{\theta}^P_{B'}) - c\widehat{\theta}^P_{B'}(\bar{\theta} + 2\widehat{\theta}^P_{B'}(\bar{\theta} - \widehat{\theta}^P_{B'}))}{2\bar{\theta}^3} > 0, \tag{A.167}$$

where the inequality is by $\bar{\theta} > 2\widehat{\theta}^P_{B'}$ by assumption.

The results (A.143), (A.164) and (A.165) imply that if the condition on F in (A.164) holds (in Case 3), ψ is U-shaped with respect to κ for $\kappa \in \left(\max\left\{\frac{4F-\bar{\theta}(P-c)}{F}, 0\right\}, 1\right)$ (which can be shown to be nonempty under (A.164)). In this case, there exists $\kappa' \in \left(\max\left\{\frac{4F-\bar{\theta}(P-c)}{F}, 0\right\}, 1\right)$ such that ψ is downward sloping with respect to κ if and only if $\kappa \in \left(\max\left\{\frac{4F-\bar{\theta}(P-c)}{F}, 0\right\}, \kappa'\right)$. If condition (A.164) is violated, then (A.143) and (A.165) imply that ψ is always upward sloping with respect to κ for $\kappa \in \left(\max\left\{\frac{4F-\bar{\theta}(P-c)}{F}, 0\right\}, 1\right)$ (if it is nonempty).

Finally, note that ψ is continuous in κ by Lemma A.4. This together with the results of Cases 1-3 implies that ψ strictly decreases in κ on some interval K if and only if any $\kappa \in K$ satisfies conditions given in Claim 1.

Claim 2. *In the pooling equilibrium, there exists threshold $\kappa^* \in \left(\frac{4F-\bar{\theta}(P-c)}{F}, 1\right)$ such that $E[U^r]$ strictly decreases in κ on some interval K if and only if for any $\kappa \in K$ it holds:*

- $\frac{\bar{\theta}(P-c)}{4} < F < \frac{1}{6}\bar{\theta}(-2c + P + \sqrt{P^2 + 2c(P-c)})$,
- $\kappa \in \left(\frac{4F-\bar{\theta}(P-c)}{F}, \kappa^*\right)$.

Proof. By Lemma A.17 we have

$$E[U^r] = \max\{\psi, 0.5(P+c)\}. \tag{A.168}$$

Hence, given continuity of ψ by Lemma A.13, $E[U^r]$ strictly decreases in κ on some interval K if and only if ψ strictly decreases in κ on this interval while $\psi > 0.5(P+c)$ for any $\kappa \in K$. By Claim 1, such interval exists if and only if

$$\frac{\bar{\theta}(P-c)^2}{2(3P+c)} < F < \frac{1}{6}\bar{\theta}(-2c+P+\sqrt{P^2+2c(P-c)}) \tag{A.169}$$

while

$$\lim_{\kappa \to \underline{\kappa}^+} \psi > 0.5(P+c) \tag{A.170}$$

where

$$\underline{\kappa} = \max\left\{\frac{4F-\bar{\theta}(P-c)}{F}, 0\right\}. \tag{A.171}$$

Let us show the necessary and sufficient conditions for (A.170). Consider the following two cases depending on whether $\underline{\kappa} = 0$.

Case 1. $\underline{\kappa} = 0$.

Note that in this case for any $\kappa \in (0,1)$ it holds

$$\kappa > 0 \geq \frac{4F-\bar{\theta}(P-c)}{F} \Rightarrow F < \bar{\theta}\frac{P-c}{4-\kappa}.$$

Then, by Lemma A.8 we have $\widehat{\theta}^P_{B'} \in (0, 0.5\bar{\theta})$ and $\widehat{\theta}^P_{N'} = 2\widehat{\theta}^P_{B'}$ for any $\kappa \in (0,1)$. Moreover, $\underline{\kappa} = 0 \Leftrightarrow \frac{4F-\bar{\theta}(P-c)}{F} \leq 0$ implies

$$F \leq \bar{\theta}\frac{(P-c)}{4}. \tag{A.172}$$

Note that

$$\lim_{\kappa\to 0} \widehat{\theta}^P_{B'} = \frac{F}{\lim_{\kappa\to 0} \eta(\bar{m}|\widehat{\theta}^P_{B'})(P-c)} = \frac{F}{0.5(P-c)} \leq \frac{\bar{\theta}\frac{P-c}{4}}{0.5(P-c)} = 0.5\bar{\theta}, \tag{A.173}$$

where the first equality is by (A.92), the second equality is by (A.158), and the inequality is by (A.172). Then, given that $\widehat{\theta}^P_{N'} = 2\widehat{\theta}^P_{B'}$ in the considered case, we have

$$\begin{aligned}\lim_{\kappa \to \underline{\kappa}^+} \psi &= \lim_{\kappa\to 0}\left(\kappa(0.5P+0.5c\frac{\widehat{\theta}^P_{B'}}{\bar{\theta}}) + (1-\kappa)\frac{2\widehat{\theta}^P_{B'}}{\bar{\theta}}(0.5P+0.5c)\right) \\ &= \frac{2\lim_{\kappa\to 0}\widehat{\theta}^P_{B'}}{\bar{\theta}}(0.5P+0.5c) \leq 0.5P+0.5c, \end{aligned} \tag{A.174}$$

where the inequality is due to $\lim_{\kappa\to 0}\widehat{\theta}^P_{B'} \leq 0.5\bar{\theta}$ by (A.173).

Hence, the condition (A.170) is not satisfied if $\underline{\kappa} = 0$.

Case 2. $\underline{\kappa} \neq 0 \Leftrightarrow \underline{\kappa} = (4F - \bar{\theta}(P-c))/F > 0$.

Note that in the considered case for any $\kappa > \underline{\kappa}$ we have

$$\kappa > \frac{4F-\bar{\theta}(P-c)}{F} \Leftrightarrow F < \bar{\theta}\frac{P-c}{4-\kappa},$$

so that by Lemma A.8 we again have $\widehat{\theta}^P_{B'} \in (0, 0.5\bar{\theta})$ and $\widehat{\theta}^P_{N'} = 2\widehat{\theta}^P_{B'}$. Consequently, by (A.99)

$$\eta(\varnothing|\widehat{\theta}^P_{B'}) = \frac{(1-\kappa)(\bar{\theta} - 2\widehat{\theta}^P_{B'})}{(2-\kappa)\bar{\theta} - (4-3\kappa)\widehat{\theta}^P_{B'}}. \tag{A.175}$$

At the same time, in the considered case

$$\lim_{\kappa \to \underline{\kappa}^+} \widehat{\theta}^P_{B'} = \widehat{\theta}^P_{B'|\kappa=\underline{\kappa}} = 0.5\bar{\theta}, \tag{A.176}$$

where the first equality is by the continuity of $\widehat{\theta}^P_{B'}$ in κ on $(0,1)$ by Lemma A.13, and the second equality is by Lemma A.4 and the fact that $\varpi(0.5\bar{\theta}) = 0$ for $\kappa = (4F - \bar{\theta}(P-c))/F$ (see (A.43)). Then, we obtain

$$\begin{aligned}\lim_{\kappa \to \underline{\kappa}^+} \eta(\varnothing|\widehat{\theta}^P_{B'}) &= \frac{(1-\underline{\kappa})(\bar{\theta} - 2\lim_{\kappa\to\underline{\kappa}} \widehat{\theta}^P_{B'})}{(2-\underline{\kappa})\bar{\theta} - (4-3\underline{\kappa})\lim_{\kappa\to\underline{\kappa}} \widehat{\theta}^P_{B'}} \\ &= \frac{(1-\underline{\kappa})(\bar{\theta} - \bar{\theta})}{(2-\underline{\kappa})\bar{\theta} - (4-3\underline{\kappa})0.5\bar{\theta}} = 0 < \underline{\eta},\end{aligned}$$

where the first equality is by (A.175) and the fact that the denominator is not equal to 0 since $\underline{\kappa} > 0$ by assumption, and the second equality is by (A.176). Consequently, for κ sufficiently close to $\underline{\kappa}$ it should also hold $\eta(\varnothing|\widehat{\theta}^P_{B'}) < \underline{\eta}$ (given that $\widehat{\theta}^P_{B'}$, and hence $\eta(\varnothing|\widehat{\theta}^P_{B'})$, is continuous in κ by Lemma A.13), i.e. the receiver will not invest conditional on $\varnothing$ so that the equilibrium is responsive. For such values of κ it would then hold $\psi > 0.5(P+c)$ by Lemma A.15(iii) and (A.168). Hence, the condition (A.170) is always satisfied if $\underline{\kappa} = (4F - \bar{\theta}(P-c))/F$.

Combining Cases 1 and 2 together implies that (A.170) is satisfied if and only if $\underline{\kappa} = (4F - \bar{\theta}(P-c))/F$, which is equivalent to

$$F > 0.25\bar{\theta}(P-c). \tag{A.177}$$

In such case, by Claim 1 and (A.170), under condition (A.164) there would exist a nonempty interval $\left(\frac{4F-\bar{\theta}(P-c)}{F}, \kappa^*\right)$ with some $\kappa^* \leq \kappa'$ such that ψ strictly decreases in κ while $\psi > 0.5(P+c)$ for any κ on this interval (and these both conditions would never hold simultaneously for any other κ). This would in turn imply by (A.168) that $E[U^r]$ strictly decreases in κ on some interval K if and only if for any $\kappa \in K$ it holds $\kappa \in \left(\frac{4F-\bar{\theta}(P-c)}{F}, \kappa^*\right)$ under the additional restrictions (A.177) and (A.164). Combining these two restrictions together results in

$$F \in \left(0.25\bar{\theta}(P-c), \frac{1}{6}(-2c\bar{\theta} + P\bar{\theta} + \bar{\theta}\sqrt{P^2 + 2cP - 2c^2})\right), \tag{A.178}$$

given that $\frac{\bar{\theta}(P-c)^2}{2(3P+c)} < 0.25\bar{\theta}(P-c)$ if the interval in (A.164) is nonempty. Note that there exist parameter values such that the interval in (A.178) is nonempty, in particular this is the case if and only if $P > -3c$. This leads to part (i) of the proposition.

(ii) *The separating equilibrium.*

Claim 1. *In the separating equilibrium, ψ strictly decreases in κ on some interval K if and only if for any $\kappa \in K$ it holds:*

- $F \in \left(\max\left\{\frac{(P-c)(1-\kappa)\bar{\theta}}{1+\sqrt{1+\kappa^2}}, \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}\right\}, \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}\right)$ for $\kappa \in \left(0, \frac{P+c}{P}\right)$,
- $F \in \left(\max\left\{\frac{(P-c)(1-\kappa)\bar{\theta}}{1+\sqrt{1+\kappa^2}}, \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}\right\}, \frac{\bar{\theta}(1-\kappa)(P+c)}{\kappa}\right)$ for $\kappa \in \left[\frac{P+c}{P}, \frac{2(P+c)}{2P+c}\right)$.

Case 1: $\widehat{\theta}^l_{N'} = \bar{\theta}$ and $\widehat{\theta}^l_{B'} = \bar{\theta}$ so that $F \geq \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}$ (by Lemma A.11). Then, by (A.134)

$$\begin{aligned} \psi &= \kappa(0.5P + 0.5c) + (1-\kappa)(0.5P + 0.5c) \\ &= 0.5P + 0.5c. \end{aligned} \tag{A.179}$$

Hence, in this case

$$\frac{\partial \psi}{\partial \kappa} = 0. \tag{A.180}$$

Case 2: $\widehat{\theta}^S_{N'} = \bar{\theta}$ and $\widehat{\theta}^S_{B'} \in (0.5\bar{\theta}, \bar{\theta})$ so that $F \in \left(\frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}, \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}\right)$ (by Lemma A.11). Then, by (A.134)

$$\psi = \kappa(0.5P + 0.5c\frac{\widehat{\theta}^S_{B'}}{\bar{\theta}}) + (1-\kappa)(0.5P + 0.5c). \tag{A.181}$$

We have

$$\frac{d\psi}{d\kappa} = \frac{\partial \psi}{\partial \kappa} + \frac{\partial \psi}{\partial \widehat{\theta}^S_{B'}}\left(\widehat{\theta}^S_{B'}\right)'. \tag{A.182}$$

At the same time,

$$\frac{\partial \psi}{\partial \kappa} = \frac{-c(\bar{\theta} - \widehat{\theta}^S_{B'})}{2\bar{\theta}}, \tag{A.183}$$

$$\frac{\partial \psi}{\partial \widehat{\theta}^S_{B'}} = \frac{c\kappa}{2\bar{\theta}}. \tag{A.184}$$

As in the previous case, by the implicit function theorem

$$\left(\widehat{\theta}^S_{B'}\right)' = -\frac{\partial \phi(\widehat{\theta}^S_{B'})/\partial \kappa}{\partial \phi(\widehat{\theta}^S_{B'})/\partial \widehat{\theta}^S_{B'}} = \frac{\left(\widehat{\theta}^S_{B'}\right)^2}{2(1-\kappa)^2\bar{\theta}}, \tag{A.185}$$

where the last equality is obtained substituting (A.62) for $\eta(m_{N'}|\widehat{\theta}^S_{B'})$ in the expression for $\phi(\widehat{\theta}^S_{B'})$. Substituting (A.183), (A.184) and (A.185) into (A.182) we get

$$\frac{d\psi}{d\kappa} = \frac{-c(\bar{\theta} - \widehat{\theta}^S_{B'})}{2\bar{\theta}} + \frac{c\kappa}{2\bar{\theta}} \frac{\left(\widehat{\theta}^S_{B'}\right)^2}{2(1-\kappa)^2\bar{\theta}}. \tag{A.186}$$

Let us find $\widehat{\theta}^S_{B'}$. Solving the indifference condition

$$\phi(\widehat{\theta}^S_{B'}) = F - \widehat{\theta}^S_{B'}\eta(m_{N'}|\widehat{\theta}^S_{B'})(P-c) = F - \widehat{\theta}^S_{B'}\frac{\bar{\theta}(1-\kappa)}{\kappa\widehat{\theta}^S_{B'} + 2\bar{\theta}(1-\kappa)}(P-c) = 0 \tag{A.187}$$

yields

$$\widehat{\theta}^S_{B'} = \frac{2F(1-\kappa)\bar{\theta}}{(P-c)(1-\kappa)\bar{\theta} - F\kappa}. \tag{A.188}$$

Substituting this into (A.186) and simplifying we obtain

$$\frac{d\psi}{d\kappa} = \frac{-c}{2(F\kappa - (P-c)(1-\kappa)\bar{\theta})^2}(a_1\bar{\theta}^2 + a_2\bar{\theta} + a_3), \tag{A.189}$$

where

$$\begin{aligned} a_1 &= ((P-c)(1-\kappa))^2, \\ a_2 &= -2(P-c)F(1-\kappa), \\ a_3 &= -F^2\kappa^2. \end{aligned}$$

Since the fraction in (A.189) is strictly positive, $\frac{d\psi}{d\kappa} < 0$ if and only if $a_1\bar{\theta}^2 + a_2\bar{\theta} + a_3 < 0$. The only positive real root of the corresponding quadratic equation is

$$\bar{\theta} = \frac{F(1+\sqrt{1+\kappa^2})}{(P-c)(1-\kappa)}. \tag{A.190}$$

Consequently, $\frac{d\psi}{d\kappa} < 0$ if

$$\bar{\theta} < \frac{F(1+\sqrt{1+\kappa^2})}{(P-c)(1-\kappa)} \Leftrightarrow F > \frac{(P-c)(1-\kappa)\bar{\theta}}{1+\sqrt{1+\kappa^2}}. \tag{A.191}$$

One can show that the RHS of (A.191) is smaller than the upper bound of F in Case 2:

$$\frac{(P-c)(1-\kappa)\bar{\theta}}{1+\sqrt{1+\kappa^2}} < \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}. \tag{A.192}$$

At the same time, it can be both larger and smaller than the lower bound $\frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}$

depending on the parameters. Consequently, $\frac{d\psi}{d\kappa} < 0$ in Case 2 if and only if

$$F \in \left(\max\left\{ \frac{(P-c)(1-\kappa)\bar{\theta}}{1+\sqrt{1+\kappa^2}}, \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa} \right\}, \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa} \right). \tag{A.193}$$

Case 3: $\bar{\theta} > \widehat{\theta}^S_{N'} = 2\widehat{\theta}^S_{B'}$ so that $F < \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}$ (by Lemma A.11). Then, by (A.134)

$$\psi = \kappa(0.5P + 0.5c\frac{\widehat{\theta}^S_{B'}}{\bar{\theta}}) + (1-\kappa)\frac{2\widehat{\theta}^S_{B'}}{\bar{\theta}}(0.5P + 0.5c). \tag{A.194}$$

We have

$$\frac{d\psi}{d\kappa} = \frac{\partial\psi}{\partial\kappa} + \frac{\partial\psi}{\partial\widehat{\theta}^S_{B'}}\left(\widehat{\theta}^S_{B'}\right)'. \tag{A.195}$$

At the same time,

$$\frac{\partial\psi}{\partial\kappa} = \frac{P(\bar{\theta} - 2\widehat{\theta}^S_{B'}) - c\widehat{\theta}^S_{B'}}{2\bar{\theta}} > 0, \tag{A.196}$$

$$\frac{\partial\psi}{\partial\widehat{\theta}^S_{B'}} = \frac{c(2-\kappa) + 2P(1-\kappa)}{2\bar{\theta}} \geq 0, \tag{A.197}$$

with the latter inequality by $\kappa \leq \frac{2(P+c)}{2P+c}$ (a necessary condition for the separating equilibrium by Proposition 4). Finally, by the implicit function theorem and the fact that $\phi(\widehat{\theta}^S_{B'}) = 0$ by Lemma A.9,

$$\left(\widehat{\theta}^S_{B'}\right)' = -\frac{\partial\phi(\widehat{\theta}^S_{B'})/\partial\kappa}{\partial\phi(\widehat{\theta}^S_{B'})/\partial\widehat{\theta}^S_{B'}}. \tag{A.198}$$

Substituting for $\eta(m_{N'}|\widehat{\theta}^S_{B'})$ in $\phi(\widehat{\theta}^S_{B'})$ (see (A.77)) and simplifying we get

$$-\frac{\partial\phi(\widehat{\theta}^S_{B'})/\partial\kappa}{\partial\phi(\widehat{\theta}^S_{B'})/\partial\widehat{\theta}^S_{B'}} = \frac{\widehat{\theta}^S_{B'}}{4(1-\kappa)^2 + \kappa(1-\kappa)} > 0. \tag{A.199}$$

(A.195)-(A.199) imply that in Case 3

$$\frac{d\psi}{d\kappa} > 0. \tag{A.200}$$

Finally, note that ψ is continuous in κ by Lemma A.4. This together with the results of Cases 1-3 implies that ψ strictly decreases in κ on some interval K if and only if for any $\kappa \in K$ it holds

$$F \in \left(\max\left\{ \frac{(P-c)(1-\kappa)\bar{\theta}}{1+\sqrt{1+\kappa^2}}, \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa} \right\}, \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa} \right). \tag{A.201}$$

Combining this condition with the existence condition for the separating equilibrium from Proposition 4, we obtain that ψ strictly decreases in κ on some interval K (while the separating equilibrium exists) if and only if for any $\kappa \in K$ it holds

$$\begin{aligned}
F &\in \left(\max\left\{\frac{(P-c)(1-\kappa)\bar{\theta}}{1+\sqrt{1+\kappa^2}}, \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}\right\}, \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}\right) \text{ for } \kappa \in \left(0, \frac{P+c}{P}\right), \\
F &\in \left(\max\left\{\frac{(P-c)(1-\kappa)\bar{\theta}}{1+\sqrt{1+\kappa^2}}, \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}\right\}, \frac{\bar{\theta}(1-\kappa)(P+c)}{\kappa}\right) \text{ for } \kappa \in \left[\frac{P+c}{P}, \frac{2(P+c)}{2P+c}\right).
\end{aligned}$$

Note that the first interval of F is always nonempty, while the second interval of F is nonempty for $\kappa \in \left[\frac{P+c}{P}, \frac{2(P+c)}{2P+c}\right)$, in particular, whenever $-c$ is sufficiently small relative to P. Altogether, this leads to Claim 1.

Claim 2. *In the separating equilibrium, $E[U^r]$ strictly decreases in κ on some interval K if and only if ψ strictly decreases in κ* on K .

Proof. Let us first show necessity. By Lemma A.17 we have

$$E[U^r] = \max\{\psi, 0.5(P+c)\}. \tag{A.202}$$

Besides, ψ is continuous in κ by Lemma A.4. Hence, $E[U^r]$ strictly decreases in κ on some interval K only if ψ strictly decreases in κ on K.

Let us show sufficiency. Assume ψ strictly decreases in κ on some interval K. By the proof of Claim 1 this can only be the case if condition (A.193) holds for any $\kappa \in K$, which implies by Lemma A.15(ii) that the equilibrium is responsive for any $\kappa \in K$. Then, by Lemma A.15(iii) it holds $E[U^r] > 0.5(P+c)$ for any $\kappa \in K$, which implies by (A.202) that $E[U^r] = \psi$ for any $\kappa \in K$. Consequently, since ψ strictly decreases in κ on K, $E[U^r]$ should decrease in κ on this interval as well.

Claims 1 and 2 together lead to part (ii) of the proposition. ■

Proof of Proposition 9. Note that generally, the expected utility of type θ before his information state is realized is, by the law of total probability,

$$\begin{aligned}
E[U^s(\theta)] &= \sum_{i^s \in \{G', N', B'\}} (U^s_{i^s}(\theta, \eta(\bar{m}), I) \Pr[\bar{m}|i^s, \theta] \Pr[i^s]) \\
&= 0.5\kappa F + 0.5\kappa \Pr[\bar{m}|B', \theta](F - \theta\eta(\bar{m})(P-c)) \\
&\quad + (1-\kappa)\Pr[\bar{m}|N', \theta](F - 0.5\theta\eta(\bar{m})(P-c)).
\end{aligned} \tag{A.203}$$

Next, consider separately the pooling and the separating equilibrium.

(i) *The pooling equilibrium.*

Claim 1. *In the pooling equilibrium, if $F \geq 0.5\bar{\theta}(P-c)$, then the sender is indifferent over his probability of being informed.*

Proof. If $F \geq 0.5\bar{\theta}(P-c)$, then by Lemma A.8 we have $\widehat{\theta}^P_{B'} = \widehat{\theta}^P_{N'} = \bar{\theta}$, so that all types send message $\bar{m}$. At the same time, it is easy to verify that $\eta(\bar{m})_{|\widehat{\theta}^P_{B'}=\widehat{\theta}^P_{N'}=\bar{\theta}} = 0.5$. After substituting this together with $\Pr[\bar{m}|B', \theta] = \Pr[\bar{m}|N', \theta] = 1$ into (A.203) we obtain

$$E[U^s(\theta)] = F - 0.25\theta(P-c),$$

which does not depend on κ. Hence, the claim follows.

Claim 2. *If $F < 0.5\bar{\theta}(P-c)$, then $\widehat{\theta}^P_{B'}$ strictly decreases with κ.*

Proof. If $F < 0.5\bar{\theta}(P-c)$, we have $\widehat{\theta}^P_{B'} < \bar{\theta}$ by Lemma A.8. Then, by (A.141) and (A.147) $\partial\widehat{\theta}^P_{B'}/\partial\kappa < 0$ (where $\widehat{\theta}^P_{B'}$ is differentiable). Consequently, since $\widehat{\theta}^P_{B'}$ is continuous in κ by Lemma A.13 (and differentiable everywhere except for a finite set of kink points), $\widehat{\theta}^P_{B'}$ strictly decreases with κ.

Claim 3. *If $F < 0.5\bar{\theta}(P-c)$, then $\eta(\bar{m}|\widehat{\theta}^P_{B'})$ strictly increases with κ.*

Proof. If $F < 0.5\bar{\theta}(P-c)$, we have $\widehat{\theta}^P_{B'} < \bar{\theta}$ by Lemma A.8. Then, by Corollary A.1 it holds

$$\eta(\bar{m}|\widehat{\theta}^P_{B'}) = \frac{F}{\widehat{\theta}^P_{B'}(P-c)}. \tag{A.204}$$

This together with Claim 2 implies that $\eta(\bar{m}|\widehat{\theta}^P_{B'})$ strictly increases with κ.

Claim 4. *In the pooling equilibrium, if $F < 0.5\bar{\theta}(P-c)$, then the sender's expected utility strictly decreases with κ on $(0,1)$ if $\theta < \underline{\theta}$ where $\underline{\theta} > 0$ is some threshold independent of κ.*

Proof. Note that $F < 0.5\bar{\theta}(P-c)$ implies $\widehat{\theta}^P_{B'}(\kappa) < \widehat{\theta}^P_{N'}(\kappa)$ for any $\kappa \in (0,1)$ by Lemma A.8. Assume further that the sender's type θ is sufficiently low so that $\theta < \widehat{\theta}^P_{B'}(\kappa) < \widehat{\theta}^P_{N'}(\kappa)$ holds for any $\kappa \in (0,1)$ (note that $\widehat{\theta}^P_{B'}(\kappa)$ is bounded from 0 for any κ by Corollary A.1). In turn, this implies that the sender sends message $\bar{m}$ in all information states so that $\Pr[\bar{m}|B', \theta] = \Pr[\bar{m}|N', \theta] = 1$. After substituting this into (A.203) we obtain

$$E[U^s(\theta)] = F - 0.5\theta\eta(\bar{m})(P-c) \tag{A.205}$$

for any $\kappa \in (0,1)$. Then, given that the equilibrium level of $\eta(\bar{m})$ strictly increases with κ by Claim 3, the claim follows.

Claim 5. *If $F < 0.5\bar{\theta}(P-c)$, then $\widehat{\theta}^P_{B'}(\kappa)$ is bounded from $\bar{\theta}$ for any $\kappa \in (0,1)$.*

Proof. From (A.91) we obtain $\lim_{\kappa\to 0}\eta(\bar{m}) = 0.5$ (given that the limit of the denominator is not 0 since $\widehat{\theta}^P_{B'}$ is bounded from 0). Then, by Corollary A.1

$$\lim_{\kappa\to 0}\widehat{\theta}^P_{B'}(\kappa) = \frac{F}{\lim_{\kappa\to 0}\eta(\bar{m})(P-c)} = \frac{F}{0.5(P-c)} < \bar{\theta}, \tag{A.206}$$

where the inequality is due to $F < 0.5\bar{\theta}(P-c)$ by assumption. From (A.206) and the fact that $\widehat{\theta}^P_{B'}(\kappa)$ is decreasing in κ by Claim 2 it follows that $\widehat{\theta}^P_{B'}(\kappa)$ is bounded from $\bar{\theta}$ for any $\kappa \in (0,1)$.

Claim 6. *In the pooling equilibrium, if* $F < 0.5\bar{\theta}(P-c)$*, then the sender's expected utility strictly increases with* κ *on* $(0,1)$ *for any* $\theta > \widetilde{\theta}$ *, where* $\widetilde{\theta} < \bar{\theta}$ *is some threshold independent of* κ*.*

Proof. Assume that the sender's type θ is sufficiently high so that $\widehat{\theta}^P_{B'}(\kappa) < \theta < \bar{\theta}$ holds for any $\kappa \in (0,1)$ (which is possible by Claim 5). This implies that the sender does not send message $\bar{m}$ in state B' so that $\Pr[\bar{m}|B',\theta] = 0$. Note that $\Pr[\bar{m}|N',\theta] = 1$, i.e. the sender induces investment while being uninformed, if and only if his expected utility from investment in state N' is positive, i.e. if and only if $F - 0.5\theta\eta(\bar{m})(P-c) \geq 0$. Denote the sender's expected utility conditional on investment in state N' as

$$\upsilon_{N'}(\theta) = U^s_{N'}(\widehat{\theta}^P_{B'}, \eta(\bar{m}|\widehat{\theta}^P_{B'}), I) = F - 0.5\theta\eta(\bar{m}|\widehat{\theta}^P_{B'})(P-c). \tag{A.207}$$

Then, (A.203) implies

$$E[U^s(\theta)] = \begin{cases} 0.5\kappa F + (1-\kappa)\upsilon_{N'}(\theta) \text{ if } \upsilon_{N'}(\theta) \geq 0, \\ 0.5\kappa F \text{ if } \upsilon_{N'}(\theta) < 0. \end{cases} \tag{A.208}$$

Differentiating with respect to κ and simplifying we get

$$\frac{\partial E[U^s(\theta)]}{\partial \kappa} = \begin{cases} 0.5\left(-\upsilon_{B'}(\theta) - (1-\kappa)\theta(P-c)\frac{\partial \eta(\bar{m})}{\partial \kappa}\right) \text{ if } \upsilon_{N'}(\theta) \geq 0, \\ 0.5F \text{ if } \upsilon_{N'}(\theta) < 0. \end{cases} \tag{A.209}$$

Clearly, $E[U^s(\theta)]$ is strictly increasing in κ on $(0,1)$ in the second case. Let us also show that $E[U^s(\theta)]$ is increasing in κ for $\kappa \in (0,1)$ in the first case, i.e. if $\upsilon_{N'}(\theta) \geq 0$, if θ is sufficiently high. Note that in this case we must have $\theta \leq \widehat{\theta}^P_{N'}$ (since type θ invests in state N') so that

$$\widehat{\theta}^P_{B'} \geq 0.5\widehat{\theta}^P_{N'} \geq 0.5\theta, \tag{A.210}$$

where the first inequality is by Lemma A.5. Next, by (A.91),

$$\eta(\bar{m}|\widehat{\theta}^P_{B'}) = \begin{cases} \frac{2\widehat{\theta}^P_{B'}(1-\kappa)+\kappa\bar{\theta}}{\widehat{\theta}^P_{B'}(4-3\kappa)+\kappa\bar{\theta}} \text{ if } \kappa < \frac{4F-\bar{\theta}(P-c)}{F}, \\ \frac{\bar{\theta}}{\kappa\widehat{\theta}^P_{B'}+(2-\kappa)\bar{\theta}} \text{ if } \kappa \geq \frac{4F-\bar{\theta}(P-c)}{F}. \end{cases} \tag{A.211}$$

Consider these two cases separately (excluding $\kappa = \frac{4F-\bar{\theta}(P-c)}{F}$, where $\eta(\bar{m}|\widehat{\theta}^P_{B'})$ is continuous but not differentiable in κ).

Case 1: $\kappa < \frac{4F-\bar{\theta}(P-c)}{F} \Leftrightarrow F < \bar{\theta}\frac{P-c}{4-\kappa}$. Then, we have $\widehat{\theta}^P_{B'} < 0.5\bar{\theta}$ by Lemma A.8. Together with (A.210), this implies

$$\widehat{\theta}^P_{B'} \in [0.5\theta, 0.5\bar{\theta}). \tag{A.212}$$

Substituting the corresponding expression for $\eta(\bar{m}|\widehat{\theta}^P_{B'})$ from (A.211) into (A.208) (under assumption $\upsilon_{N'}(\theta) \geq 0$) and differentiating with respect to κ (while using (A.147) for $\frac{\partial \widehat{\theta}^P_{B'}}{\partial \kappa}$), and then taking the limit of the resulting expression as θ converges to $\bar{\theta}$ (and hence $\widehat{\theta}^P_{B'}$ converges to $0.5\bar{\theta}$ due to (A.212)), we obtain

$$\lim_{\theta \to \bar{\theta}} \frac{dE[U^s(\theta)]}{d\kappa} = \frac{3\bar{\theta}(P-c)}{(4-\kappa)^2} - 0.5F > 0, \tag{A.213}$$

where the inequality follows from $F < \bar{\theta}\frac{P-c}{4-\kappa}$ by assumption. Moreover, this expression is bounded from 0 for $F < \bar{\theta}\frac{P-c}{4-\kappa}$. Hence, there must exist $\theta' < \bar{\theta}$ (independent of κ) such that $\frac{dE[U^s(\theta)]}{d\kappa} > 0$ for any $\theta > \theta'$ and any κ in the considered case, unless the derivative of $\frac{dE[U^s(\theta)]}{d\kappa}$ with respect to θ may converge to positive infinity. Yet, this is excluded by (A.209) given that $\upsilon_{B'}(\theta)$ is bounded and $\frac{\partial \eta(\bar{m})}{\partial \kappa} > 0$ by Claim 3.

Case 2: $\kappa > \frac{4F - \bar{\theta}(P-c)}{F} \implies F \in \left(\bar{\theta}\frac{P-c}{4-\kappa}, 0.5\bar{\theta}(P-c)\right)$ (given the initial restriction of Claim 6).

Then, we have $\widehat{\theta}^P_{B'} \in (0.5\bar{\theta}, \bar{\theta})$ by Lemma A.8. Note that in this case the equilibrium condition $\varpi(\widehat{\theta}^P_{B'}) = 0$ yields a closed-form solution for $\widehat{\theta}^P_{B'}$:

$$\widehat{\theta}^P_{B'} = \frac{F\bar{\theta}(2-\kappa)}{\bar{\theta}(P-c) - F\kappa}. \tag{A.214}$$

Substituting this together with the corresponding expression for $\eta(\bar{m}|\widehat{\theta}^P_{B'})$ from (A.211) into (A.208) (under assumption $\upsilon_{N'}(\theta) \geq 0$) and differentiating with respect to κ, and then taking the limit of the resulting expression as θ converges to $\bar{\theta}$, we obtain

$$\lim_{\theta \to \bar{\theta}} \frac{dE[U^s(\theta)]}{d\kappa} = \frac{(P-c)^2\bar{\theta}^2 - 2F^2(1-\kappa)\kappa - F(2-\kappa+\kappa^2)(P-c)\bar{\theta}}{2(2-\kappa)^2(P-c)\bar{\theta}},$$

which can be shown to be always strictly positive for $F < 0.5\bar{\theta}(P-c)$. Moreover, by the same argument as in Case 1, there must exist $\theta'' < \bar{\theta}$ (independent of κ) such that $\frac{dE[U^s(\theta)]}{d\kappa} > 0$ for any $\theta > \theta''$ and any κ in the considered case.

In sum, Cases 1 and 2 imply that there exists $\widetilde{\theta} < \bar{\theta}$ independent of κ such that $\frac{dE[U^s(\theta)]}{d\kappa} > 0$ (where it exists) for any $\theta > \widetilde{\theta}$ and any κ (as far as $\upsilon_{N'}(\theta) \geq 0$ still holds). Given that $\frac{dE[U^s(\theta)]}{d\kappa} > 0$ for any $\kappa \in (0,1)$ as far as $\upsilon_{N'}(\theta) < 0$ by (A.209), we obtain that $\frac{dE[U^s(\theta)]}{d\kappa} > 0$ for any $\theta > \widetilde{\theta}$ and any $\kappa \in (0,1)$ whenever this derivative exists.

Finally, note that $\eta(\bar{m})$ is continuous in κ by (A.204) and the fact that $\widehat{\theta}^P_{B'}$ is continuous in κ by Lemma A.13. Consequently, by (A.208), $E[U^s(\theta)]$ is also continuous in κ. Given that $E[U^s(\theta)]$ is also differentiable in κ except on a finite set of kink points, while $\frac{dE[U^s(\theta)]}{d\kappa} > 0$ for any $\theta > \widetilde{\theta}$ and $\kappa \in (0,1)$ where differentiable (as shown above), Claim 6 follows.

Part (i) of the proposition jointly follows from Claims 1, 4 and 6.

(ii) *The separating equilibrium.*

Claim 7. *$\widehat{\theta}^S_{B'}$ weakly increases with κ, and strictly so if and only if $\widehat{\theta}^S_{B'} < \bar{\theta}$.*

Proof. Consider the separating equilibrium under different values of F. If $F \geq \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}$, then by Lemma A.11 $\widehat{\theta}^S_{B'} = \bar{\theta}$ so that $\partial\widehat{\theta}^S_{B'}/\partial\kappa = 0$. If $F < \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa} \Longrightarrow \widehat{\theta}^S_{B'} < \bar{\theta}$ (by Lemma A.8), then by (A.185) and (A.199) $\partial\widehat{\theta}^S_{B'}/\partial\kappa > 0$ (where $\widehat{\theta}^S_{B'}$ is differentiable). Consequently, since $\widehat{\theta}^S_{B'}$ is continuous in κ by Lemma A.13 (and differentiable everywhere except on a finite set of kink points), $\widehat{\theta}^S_{B'}$ weakly increases with κ, and strictly so if and only if $\widehat{\theta}^S_{B'} < \bar{\theta}$.

Claim 8. *$\eta(\widetilde{m}|\widehat{\theta}^S_{B'})$ is continuous in κ.*

Proof. By Lemma 5, $\eta(\widetilde{m}|\widehat{\theta}^S_{B'})$ is generally given by

$$\begin{aligned}\eta(\widetilde{m}|\widehat{\theta}^S_{B'}) &= \frac{\Pr[\widetilde{m}|G']\kappa + \Pr[\widetilde{m}|N'](1-\kappa)}{(\Pr[\widetilde{m}|G'] + \Pr[\widetilde{m}|B'])\kappa + 2\Pr[\widetilde{m}|N'](1-\kappa)} \\ &= \frac{\widehat{\theta}^S_{N'}(1-\kappa)}{\widehat{\theta}^S_{B'}\kappa + 2\widehat{\theta}^S_{N'}(1-\kappa)} \\ &= \frac{\min\left\{\bar{\theta}, 2\widehat{\theta}^S_{B'}\right\}(1-\kappa)}{\widehat{\theta}^S_{B'}\kappa + 2\min\left\{\bar{\theta}, 2\widehat{\theta}^S_{B'}\right\}(1-\kappa)},\end{aligned} \tag{A.215}$$

where the last equality is by Lemma A.5. Hence, $\eta(\widetilde{m}|\widehat{\theta}^S_{B'})$ is continuous in $\widehat{\theta}^S_{B'}$. At the same time, $\widehat{\theta}^S_{B'}$ is continuous in κ by Lemma A.13, which leads to the claim.

Claim 9. *$\eta(\widetilde{m}|\widehat{\theta}^S_{B'})$ strictly decreases with κ.*

Proof. Note that by Corollary A.2 it holds in the separating equilibrium

$$\eta(\widetilde{m}|\widehat{\theta}^S_{B'}) = \begin{cases} \frac{F}{\widehat{\theta}^S_{B'}(P-c)} \text{ if } \widehat{\theta}^S_{B'} < \bar{\theta}, \\ \eta(\widetilde{m}|\bar{\theta}) = \frac{1-\kappa}{2-\kappa} \text{ if } \widehat{\theta}^S_{B'} = \bar{\theta}. \end{cases} \tag{A.216}$$

Note that both functions strictly decrease with κ (with the first one by Claim 7). At the same time, $\eta(\widetilde{m}|\widehat{\theta}^S_{B'})$ must be continuous in κ by Claim 8. Note also that $\widehat{\theta}^S_{B'}$ is monotonic in κ by Claim 7. Hence, $\eta(\widetilde{m}|\widehat{\theta}^S_{B'})$ strictly decreases with κ.

Claim 10. *In the separating equilibrium, the sender's expected utility strictly increases with κ for any θ.*

Proof. Denote again the sender's expected utility conditional on investment in state $i^s \in \{B', N'\}$ as

$$\upsilon_{i^s}(\theta) = U^s_{i^s}(\widehat{\theta}^S_{B'}, \eta(\widetilde{m}|\widehat{\theta}^S_{B'}), I) = F - \Pr[B|i^s]\theta\eta(\widetilde{m}|\widehat{\theta}^S_{B'})(P-c). \tag{A.217}$$

By incentive compatibility, $\Pr[\widetilde{m}|B'] = 1$ if and only if $\upsilon_{B'}(\theta) \geq 0$, while $\Pr[\widetilde{m}|N'] = 1$ if and only if $\upsilon_{N'}(\theta) \geq 0$. Then by (A.203),

$$E[U^s(\theta)] = \begin{cases} 0.5\kappa F + 0.5\kappa\upsilon_{B'}(\theta) + (1-\kappa)\upsilon_{N'}(\theta) \text{ if } \upsilon_{B'}(\theta) \geq 0, \\ 0.5\kappa F + (1-\kappa)\upsilon_{N'}(\theta) \text{ if } \upsilon_{B'}(\theta) < 0 \wedge \upsilon_{N'}(\theta) \geq 0, \\ 0.5\kappa F \text{ if } \upsilon_{N'}(\theta) < 0. \end{cases} \tag{A.218}$$

Differentiating with respect to κ and simplifying we get

$$\frac{\partial E[U^s(\theta)]}{\partial \kappa} = \begin{cases} 0.5\kappa\frac{\partial \upsilon_{B'}(\theta)}{\partial \kappa} + (1-\kappa)\frac{\partial \upsilon_{N'}(\theta)}{\partial \kappa} \text{ if } \upsilon_{B'}(\theta) \geq 0, \\ -0.5\upsilon_{B'}(\theta) + (1-\kappa)\frac{\partial \upsilon_{N'}(\theta)}{\partial \kappa} \text{ if } \upsilon_{B'}(\theta) < 0 \wedge \upsilon_{N'}(\theta) \geq 0, \\ 0.5F \text{ if } \upsilon_{N'}(\theta) < 0. \end{cases} \tag{A.219}$$

Note that $-0.5\upsilon_{B'}(\theta)$ in the second case is positive given that $\upsilon_{B'}(\theta) < 0$ in this case. At the same time, for any $i^s \in \{B', N'\}$ we have (in cases where $\eta(\widetilde{m}|\widehat{\theta}^S_{B'})$ is differentiable in κ)

$$\frac{\partial \upsilon_{i^s}(\theta)}{\partial \kappa} = -\Pr[B|i^s]\theta(P-c)\frac{\partial \eta(\widetilde{m}|\widehat{\theta}^S_{B'})}{\partial \kappa} > 0,$$

where the inequality is by Claim 9. Thus, by (A.219), $\frac{\partial E[U^s(\theta)]}{\partial \kappa} > 0$ (where it is differentiable) in all cases. This together with the fact that $E[U^s(\theta)]$ is continuous in κ since $\eta(\widetilde{m}|\widehat{\theta}^S_{B'})$ is continuous in κ by Claim 8 (while $E[U^s(\theta)]$ is also differentiable in κ except on a finite set of kink points), implies that $E[U^s(\theta)]$ is strictly increasing in κ.

Part (ii) of the proposition follows from Claim 10. ■

Proof of Proposition 10. Consider the game under outcome-based preferences. Without loss of generality, we consider the pooling equilibrium (recall that all other equilibria are payoff-equivalent by Proposition 1).

Claim 1. *If an equilibrium is unresponsive, then the receiver obtains an expected payoff of* $0.5(P+c)$. *Any responsive equilibrium yields a strictly higher expected payoff.*

Proof. The proof follows by the same arguments as the proof of Lemma A.15(iii).

Claim 2. *In any equilibrium,* $E[U^r] = \max\{\psi, 0.5(P+c)\}$, *where*

$$\psi = \kappa(0.5P + 0.5c\frac{\widehat{\theta}_{B'}}{\bar{\theta}}) + (1-\kappa)\frac{\widehat{\theta}_{N'}}{\bar{\theta}}0.5(P+c).$$

Proof. The proof follows by the same arguments as the proof of Lemma A.17 given Claim 1.

Claim 3. *ψ continuously increases in κ .*

Proof. We have

$$\begin{aligned} \psi &= \kappa(0.5P + 0.5c\frac{\widehat{\theta}_{B'}}{\bar{\theta}}) + (1-\kappa)\frac{\widehat{\theta}_{N'}}{\bar{\theta}}(0.5P+0.5c) && \text{(A.220)} \\ &= \kappa(0.5P + 0.5c\frac{\min\{\bar{\theta}, F/\rho\}}{\bar{\theta}}) + (1-\kappa)\frac{\min\{\bar{\theta}, 2F/\rho\}}{\bar{\theta}}(0.5P+0.5c), && \text{(A.221)} \end{aligned}$$

where the last equality is by (A.8) and (A.9). Then,

$$\frac{\partial \psi}{\partial \kappa} = \frac{1}{2\bar{\theta}}(P(\bar{\theta} - \min\{\bar{\theta}, 2F/\rho\}) - c(\min\{\bar{\theta}, 2F/\rho\} - \min\{\bar{\theta}, F/\rho\})) \geq 0. \tag{A.222}$$

Claim 4. *$E[U^r]$ is non-decreasing in κ .*
Proof. The result follows directly from Claims 2 and 3.

Claim 5. *$E[U^s]$ is non-decreasing in κ.*
Proof. By the law of total probability,

$$\begin{aligned} E[U^s(\theta)] &= \sum_{i^s \in \{G', N', B'\}} (U^s_{i^s}(\theta, \eta(\bar{m}), I) \Pr[\bar{m}|i^s, \theta] \Pr[i^s]) \\ &= 0.5\kappa F + 0.5\kappa \Pr[\bar{m}|B', \theta](F - \theta\rho) \\ &\quad + (1-\kappa) \Pr[\bar{m}|N', \theta](F - 0.5\theta\rho). \end{aligned} \tag{A.223}$$

Then, redefining

$$\upsilon_{i^s}(\theta) = U^s_{i^s}(\theta, I) = F - \Pr[B|i^s]\theta\rho, \tag{A.224}$$

analogously to (A.218) we obtain

$$E[U^s(\theta)] = \begin{cases} 0.5\kappa F + 0.5\kappa \upsilon_{B'}(\theta) + (1-\kappa)\upsilon_{N'}(\theta) \text{ if } \upsilon_{B'}(\theta) \geq 0, \\ 0.5\kappa F + (1-\kappa)\upsilon_{N'}(\theta) \text{ if } \upsilon_{B'}(\theta) < 0 \wedge \upsilon_{N'}(\theta) \geq 0, \\ 0.5\kappa F \text{ if } \upsilon_{N'}(\theta) < 0. \end{cases} \tag{A.225}$$

Differentiating with respect to κ and simplifying yields

$$\frac{\partial E[U^s(\theta)]}{\partial \kappa} = \begin{cases} 0 \text{ if } \upsilon_{B'}(\theta) \geq 0, \\ -0.5\upsilon_{B'}(\theta) \text{ if } \upsilon_{B'}(\theta) < 0 \wedge \upsilon_{N'}(\theta) \geq 0, \\ 0.5F \text{ if } \upsilon_{N'}(\theta) < 0. \end{cases} \tag{A.226}$$

Note that in the second case $-0.5\upsilon_{B'}(\theta) > 0$ since $\upsilon_{B'}(\theta) < 0$ in that case. Thus, $\frac{\partial E[U^s(\theta)]}{\partial \kappa} \geq 0$ in all cases.

The statement of the proposition follows from Claims 4 and 5. ■